

RESULT	12
LOCUS	BG298479
DEFINITION	BC298479 900 bp mRNA EST 21-FEB-2001
ACCESSION	602396747F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4511434 5'
VERSION	BC298479
KEYWORDS	BC298479.1 GI:13063173
SOURCE	EST.
ORGANISM	house mouse. Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
AUTHORS	1 (bases 1 to 900)
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D.

Plate: LAM10394 row: n column: 11
High quality sequence stop: 676.

```

source
1. .900
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image="IMAGE:4511434"
/clone_id="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pcwv-SPORE6; Site:1; NotI:
Site:2; SalI: Cloned unidirectionally; oligo-dt primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies
Note: this is a NIH-MGC library."

```

[illegible]

RESULT	13
LOCUS	BF607177
DEFINITION	BF607177 830 bp mRNA
ACCESSION	U11000119 Mouse cDNA library ICRP522 Mus musculus
VERSION	BF607177.1 GI:13503669

KEYWORDS	EST.
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
	house mouse.
	Mus musculus.
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus 1 (passes 1 to 830)
	Yahyawi, M., Hennig, S., Neidhardt, L., Radelof, U., Hermann, B. G., Leberich, H. and O'Brien, J.
	Detection of a high number of novel genes in a 9-day mouse embryo cDNA library normalised by oligonucleotide fingerprinting
	Unpublished (2001)
	Contact: Hennig S

EST's are made from clones being representative of clone clusters.
 Clone clusters were calculated from oligonucleotide fingerprints.
 PCR primers
 FORWARD: 5'-GAGCTATTCCAGAAGTACTGA-3',
 BACKWARD: 5'-TAATACGACTCACTATAGGG-3',
 Seq primer: 5'-ATTAGTAGCAGACACTATAG-3',
 High quality sequence stop: 830.
 Location/Qualifiers

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="ICRfp522i0930"
/clone_lib="Mouse 9-day fetus cdna library ICRfp522"
/tissue_type="whole embryo"
/dec_stage="embryonic 9-day"
/lab_host="E.coli, XL1 blue"
/node="Vector: pSVSPORT1; Site_1: NotI; Site_2: SalI;
Library preparation by oligo-dT priming of RNA. Clones can
be ordered from the Resource Center in Berlin,
http://www.rzpd.de."

```

Query Match	63.3%	Score 140.6	DB 11	Length 830	
Best Local Similarity	81.4%	Pred. No. 2e-14			
Matches 193	Conservative	0	Mismatches 29	Indels 15	Gaps 2
QY 1	gacatcagcacacaycgagaggaaaccagctcgcgcgcgtctcgtcagcagaatcaygc	60			
Db 265	GACATTAAGTGACAGCGGAGAGAAACCAGCTCAGCTGAGTTCATCATCGACAAATCAGC	324			
QY 61	tcaggtcttgaagaagaagaaggagagagaaag---gagagagagaaaggagaccc	117			
Db 325	TCAAGTCTTAAAGAGAGGAGGAGGAGGAAGAAAGAAAGAAAGAAAGAGGAGGCACG	384			
QY 118	agtgaagatcagagagagaaagaa-----gagagagagagagagacccg	165			
Db 385	AGTGAAGAAATCAATAGGAGAAAGAAAGAAAGAAAGAGGAGGAGGAAGAGAGACTGG	444			
QY 166	agcaacctctagagagatcagaacagcttcgcgaagaagaagtgtggagaatatg	222			
Db 445	AGCAACTCTAGAGAGGCGCTTGAAACAATCTCCGCAATAGCAATGCTAGTGAAGAAATG	501			

RESULT 14			
BF998861			
LOCUS	BF998861	456 bp	
DEFINITION	MR3-GN0152-101100-003-b03 GN0152	EST	24-JAN-2001
ACCESSION	BF998861	Homosapiens	CDNA, mRNA sequence
VERSION	BF998861.1	GI:12434720	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

OY	120	tgaagaatca..gagaggaaagagagagggagagagagaacccgcgcgaactcttgag	178
Dd	373	TGAATAATCAGGAGGAGAAGAGAGAGAGAGAGAGAGACCGCACCACTCTTGAAG	432
OY	179	aggcatcagagca	191
Dd	433	AGGCATCAGACGA	445
RESULT	8		
LOCUS	AM804794	283 bp	mRNA EST 16-MAY-2000
DEFINITION	OYA-U00094-300300-156-g11 U00094 Homo sapiens cDNA, mRNA sequence.		
ACCESSION	AM804794		
VERSION	AM804794..1 GI:7856768		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 283)		
AUTHORS	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Grunstein,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Goldman,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.		
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed		
JOURNAL	Sequence tags		
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)		
COMMENT	20202663		
	Contact: Stimpson A.J.G.		

```

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=4t2-qv4-UM0094-300
300-156-g11a1c3=2000-03-30&4=1)
Seq primer: puc 18 forward
High quality sequence start: 2
High quality sequence stop: 283.
Location/Qualifiers
1. 283
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_idb="UM0094"
/dev_strage="A9ult"
/note="Organ: uterus; Vector: puc18; Site_1: Sma1; Site_2
Sma1; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

```

	Query Match	75.3%;	Score 167.2;	DB 10;	Length 283;
	Best Local Similarity	91.4%;	Pred. No. 8.7e-19;		
	Matches 203;	Conservative	0;	Mismatches 8;	Indels 11;
					Gaps 2;
Oy	10	gacacgagagaaagaccagctcgagccagatctctgcagcagaatcaagctcagttct	69		
Db	44	gtcaccgagagagacgac--acgtcgagcgagtcctcgttcagcggaaftacagctcagcttct	101		
Oy	70	gaggaagaagagagagagagagaaagagagagagaaaggagacacagatgaagaatca	129		

[illegible]

RESULT	9								
LOCUS	BE742222								
DEFINITION	BE742222	723 bp	mRNA	EST	15-SEP-2000				
	601575623P1	NIH_MGC_9	Homo sapiens cDNA clone IMAGE:3836647 5' ,						
ACCESSION	BE742222								
VERSION	BE742222.1	GI:10156214							
KEYWORDS	EST.								
SOURCE	human.								
ORGANISM	Homo sapiens								
	Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.								
REFERENCE	1 (bases 1 to 723)								
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .								
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)								
JOURNAL	Unpublished (1999)								
COMMENT	Contact: Robert Strausberg, Ph.D.								

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: Rhe I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: Image.lln.gov
Plate: LINC520 row: j column: 08
High quality sequence stop: 702.

FEATURES	source	Location/Qualifiers
BASE COUNT	230 a	1..723
ORIGIN	140 c	/organism="Homo sapiens" /db_xref="taxon:9606" /clone_1="IMAGE:3836647" /clone_1b="NH_MCC_9" /rclone_type="adenocarcinoma cell line" /lab_host="DH10B (phage-resistant)" /notes="Organ: ovary; Vector: pORV7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)." 284 g 69 t

Query Match	Best Local Similarity	Score 152.4;	DB 10;	Length 723;
Matches 206;	Conservative	87.3%;	Pred No. 2.3e-16;	Mismatches 16; Indels 14; Gaps 3
QY 1	gacatcagcaacagcgaagaaagaaacacgctcggccgagtcctctcgtcacaagaataagcc	60		
Db 407	GACATCAGCGACAGCGAGAGGAAGACACTCTGGCCGATCTCCTCGTCACCGGAATCAGCG	466		
QY 61	tcaagttctcgaagaagaa---gaggaggaaggaagaaggaaggaaggaaggaagc	116		
Db 467	TCAGGTTCTGAGGAGGAAGAACGAGAGAGCGGACGACAGCAGGAGGACGAGACCGGAGCAC	526		
QY 117	cagtgaaagatca-----gaggaggaagaaggaagggaggaaggaaggaaccgag	167		
Db 527	CAGTGACGAATCAGACGACGAGGAGGAGGAAGAGGAAGAGGAGGAGGAGGACCGGACG	586		

Oy		61	tcaagttctbaagaagaaggaggaggaaggggagggagggagggagacact	120
Dd	305	TCAAGTCTCCTAGGAAAGAACAACGAACGACGGAAAAGAGCGAAGAACACCACACT	364	
Oy	121	gaagatcagaag9-----aggaaagagaagaagaagaagaccgcagacaac	171	
Dd	365	GAGGAATTCACAGGCAGGAGGAGGAGCAAAGGAGGAGGAGGAGCACCGGACGACAC	424	
Oy	172	tcgaagagagcatcacagagcgtctccgaagaagtgaagtgsagaanaag	222	
Dd	425	TCTGAGAGCATCAGACGAGTGCTCCGACAGAGTAAGTAGGCAAAAG	475	
RESULT	6	Bt108825	493 bp	mRNA EST 14-JUN-2001
LOCUS	Bt108825/c	IL3-MT0267-050101-435-H02 MT0267	Homo sapiens cDNA,	mRNA sequence.
DEFINITION	Bt108825			
ACCESSION	Bt108825.1	GI:14425455		
VERSION				
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R., Nagal,M.A.F., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher.P., Jongeneel,C.V., O'Hare .M.J., Soares,F., Brentanl,R.R., Reis,I.F., de Souza,S.J. and Simpson,A.J.C. Shotgun sequencing of the human transcriptome with ORF expressed sequence tags Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000) 20202663 Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001 Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Genome Project. This entry can be seen in the following URL http://www.ludwig.org.br/scripts/gethtml2.pl?I=II3&t2=III3-MT0267 050101-435-H02&t3=2001-01-05&t4=1) Seq primer: puc 18 forward High quality sequence start: 13 High quality sequence stop: 283. Location/Qualifiers 1..493 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="MT0267" /deposition_date="Adult"note="Organ: marrow; Vector: puc18; Site_1: Sma1; Site_2: from OKRESTES PCR (U.S. Letters Patent application No. 196 into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."			
BASE COUNT	66 a	178 C	90 g	159 t
ORIGIN				
Query Match	81.4%; Score 180.6;	DB 11;	Length 493;	
Best Local Similarity	90.0%;	Pred. No. 5,36-21;		
Matches	208; Conservative	0; Mismatches 14;	Indels 9;	Gaps 14;
Oy	1	gatcatcgacagcagcagcagcagcagctcgagccagtgcttgtcagcagaattcagc	60	
I				

QY	Db	121	QY	Db	122	QY	Db	123	QY	Db	124	QY	Db	125	QY	Db	126	QY	Db	127	QY	Db	128	QY	Db	129	QY	Db	130	QY	Db	131	QY	Db	132	QY	Db	133	QY	Db	134	QY	Db	135	QY	Db	136	QY	Db	137	QY	Db	138	QY	Db	139	QY	Db	140	QY	Db	141	QY	Db	142	QY	Db	143	QY	Db	144	QY	Db	145	QY	Db	146	QY	Db	147	QY	Db	148	QY	Db	149	QY	Db	150	QY	Db	151	QY	Db	152	QY	Db	153	QY	Db	154	QY	Db	155	QY	Db	156	QY	Db	157	QY	Db	158	QY	Db	159	QY	Db	160	QY	Db	161	QY	Db	162	QY	Db	163	QY	Db	164	QY	Db	165	QY	Db	166	QY	Db	167	QY	Db	168	QY	Db	169	QY	Db	170	QY	Db	171	QY	Db	172	QY	Db	173	QY	Db	174	QY	Db	175	QY	Db	176	QY	Db	177	QY	Db	178	QY	Db	179	QY	Db	180	QY	Db	181	QY	Db	182	QY	Db	183	QY	Db	184	QY	Db	185	QY	Db	186	QY	Db	187	QY	Db	188	QY	Db	189	QY	Db	190	QY	Db	191	QY	Db	192	QY	Db	193	QY	Db	194	QY	Db	195	QY	Db	196	QY	Db	197	QY	Db	198	QY	Db	199	QY	Db	200	QY	Db	201	QY	Db	202	QY	Db	203	QY	Db	204	QY	Db	205	QY	Db	206	QY	Db	207	QY	Db	208	QY	Db	209	QY	Db	210	QY	Db	211	QY	Db	212	QY	Db	213	QY	Db	214	QY	Db	215	QY	Db	216	QY	Db	217	QY	Db	218	QY	Db	219	QY	Db	220	QY	Db	221	QY	Db	222	QY	Db	223	QY	Db	224	QY	Db	225	QY	Db	226	QY	Db	227	QY	Db	228	QY	Db	229	QY	Db	230	QY	Db	231	QY	Db	232	QY	Db	233	QY	Db	234	QY	Db	235	QY	Db	236	QY	Db	237	QY	Db	238	QY	Db	239	QY	Db	240	QY	Db	241	QY	Db	242	QY	Db	243	QY	Db	244	QY	Db	245	QY	Db	246	QY	Db	247	QY	Db	248	QY	Db	249	QY	Db	250	QY	Db	251	QY	Db	252	QY	Db	253	QY	Db	254	QY	Db	255	QY	Db	256	QY	Db	257	QY	Db	258	QY	Db	259	QY	Db	260	QY	Db	261	QY	Db	262	QY	Db	263	QY	Db	264	QY	Db	265	QY	Db	266	QY	Db	267	QY	Db	268	QY	Db	269	QY	Db	270	QY	Db	271	QY	Db	272	QY	Db	273	QY	Db	274	QY	Db	275	QY	Db	276	QY	Db	277	QY	Db	278	QY	Db	279	QY	Db	280	QY	Db	281	QY	Db	282	QY	Db	283	QY	Db	284	QY	Db	285	QY	Db	286	QY	Db	287	QY	Db	288	QY	Db	289	QY	Db	290	QY	Db	291	QY	Db	292	QY	Db	293	QY	Db	294	QY	Db	295	QY	Db	296	QY	Db	297	QY	Db	298	QY	Db	299	QY	Db	300	QY	Db	301	QY	Db	302	QY	Db	303	QY	Db	304	QY	Db	305	QY	Db	306	QY	Db	307	QY	Db	308	QY	Db	309	QY	Db	310	QY	Db	311	QY	Db	312	QY	Db	313	QY	Db	314	QY	Db
----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----	-----	----	----

Dd	276	GAAATCAGCGACGAGGAGGAAGAACCAAGCTCGGGCCAGTCTCTGTCAAGCGAATTCAGGC	335
Oy	61	tcaaggtcttcgaggaaagaagaggaagaaagaagaagaagaagacaccagt	120
Dd	336	TCAGGTTCTTGAGGAAGAAGAGAGAGAGAGAAAGAGGAGGAGAAAGGAGACACAGT	395
Oy	121	gaagaatcata-----gaagagaaagaagaagaagaagaagaacccgcagacaac	171
Dd	396	GAAAGATCATGAGAGGAGGAGGAGGAGGAAGAGGAGGAGGAGGAGGAGACCCTGCAGCAC	455
Oy	172	tcttgagggaggcatcagaagcagtcgcgaaagaatgaatgaggaagaata	222
Dd	456	TCTGAGGAGGCAATCAGACGACTCTGCCGAAGAGTAAGTAGGAGGAAGAAATG	506
RESULT	4		
LOCUS	B1261348		
DEFINITION	B1261348	345 bp mRNA EST	17-JUL-2001
ACCESSION	602953350P1 NIH_MGC_99	Homo sapiens cDNA clone IMAGE:5087556	5'
VERSION	B1261348	mRNA sequence.	
KEYWORDS	B1261348.1	GI:14820527	
SOURCE	EST.		
ORGANISM	human.		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	NIH-MGC	http://mgs.nci.nih.gov/.	
AUTHORS	1 (bases 1 to 345)		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D.		

```

FEATURES
SOURCE
118 a      57 c      131 g      39 t
Library="
Tissue Procurement: Dr. Louis Staudt, M.D., Ph.D. CDNA Library
Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.liml.gov
Plate: LUCM1845 row: c column: 13
High quality sequence stop: 345.
Location/Qualifiers
1. 345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5087556"
/clone_lib="NIH_MGC_99"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lymph; Vector: pORF; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(C). Size selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-CDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library."

```

	Query Match	88.8%	Score 197.2;	DB 11;	Length 345;
	Best Local Similarity	94.4%;	Pred. NO. 1e-23;		
	Matches 219; Conservative	0;	Mismatches 3;	Indels 10;	Gaps 1;
Oy	1	gacatcagcgaaagaagaaagcaaccagctgcggccgaagtccctcg	tcaagcgaatcacgc	60	
Dd	13	GACATCAGCGAAGGAGGAGAAGCAACGCTGGCCGAGTCTCTCTCAGCGGAATCAGGC	72		
	61	tcaagttctgaagaagaagaagagagaaagaagaagaagaagaaggagcaacct	120		

[illegible]

RESULT		5		
BP156008				
LOCUS	BP156008	516 bp	mRNA	EST
DEFINITION	RCC-HH09055-270900-033-d08 HT0955 Homo sapiens CDNA,			30-OCT-2000
ACCESSION	BF156008			
VERSION	BF156008.1		GI:11051191	
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 516)			
	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Britones,M.R.,			
	Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,			
	Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,			
	Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare			
	,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and			
	Simpson,A.J.			
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed			
	sequence tags			
JOURNAL	Proc Natl. Acad. Sci. U.S.A.	97 (7),	3491-3496	(2000)
MEDLINE	20202663			
COMMENT	Contact: Simpson A.J.G.			

```

FEATURES
source
    Ludwig Institute for Cancer Research
    Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
    Brazil
    Tel: +55-11-2704922
    Fax: +55-11-2707001
    Email: asimpson@ludwig.org.br
    This sequence was derived from the FAPESP/LICR Human Cancer Genome
    Project. This entry can be seen in the following URL
    (http://www.ludwig.org.br/scripts/gethtml2.pl?fl1=at2-RC0-HR0955-2707)
    900-033-gd08&t3=2000-09-27&t4=1)
    Seq primer: puc 18 forward
    High quality sequence start: 56
    High quality sequence stop: 516.
    Location/Qualifiers
        1..516
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_id="HR0955"
            /dev_stage="Adult"
            /note="Organ: head/neck; Vector: puc18; Site:1: SmaI;
            Site:2: SmaI; A mini-library was made by cloning products
            derived from ORESTES PCR (U.S. Letters Patent application
            No. 196,716 - Ludwig Institute for Cancer Research)
            profiles into the pUC 18 vector. Reverse transcription of
            tissue mRNA and cDNA amplification were performed under
            low stringency conditions."
155 a      109 c      175 r      77 t

```

Query Match	88.6%	Score 196.6	DB 11	Length 516
Best Local Similarity	94.4%	Pred. No. 1.2e-23		
Matches 218	Conservative 0	Mismatches 4	Indels 9	Gaps 1
Qy	1	gacatcagcgcagacgcgagagagaaacacagctctgcgcgagctcttcgttcagcagaalcaagc	60	
Db	245	GACATCAGCGAGCAGGAGGAAAGACCACTCTGGCCGCGATCTCTGTCACCGGGAATTCAGGC	304	

```

/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      230 a      155 c      254 g      100 t      1 others
ORIGIN

```

Query Match	90.7%;	Score 201.4;	DB 10;	Length 740;
Best Local Similarity	95.7%;	Pred. No. 2e-24;		
Matches 221; Conservative	0;	Mismatches 1;	Indels 9;	Gaps 1;

[illegible]

RESULT	2				
LOCUS	BE740559				
DEFINITION	BE740559	1063 bp	mRNA	EST	15-SEP-2000
	BE740559	NIH_MGC_9	Homo sapiens	cdna clone	IMAGE:3949693 5' ,
ACCESSION	BE740559				
VERSION	BE740559.1	GI:10154551			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				

REFERENCE	1 (bases 1 to 1063)
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

Email: cgap@us-1.nhl.nih.gov
Tissue Procurement: DCTP/DTF
cDNA library preparation: Ling Hong/Rubin Laboratory
DNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
plate: LNCM614 row: p column: 14
high quality sequence stop: 741.

FEATURES

1. .1063

```

/organism="Homo sapiens"
/db_xref="taxon:3606"
/clone="IMAGE:3949603"
/clone_lib="NTH_MGC_9"
/tissue_type="adenocarcinoma cell line"
lab_host="nuovo (phorbol-treated)"

```

BASE CONTINUED

351 a 203 c 402 g 107 t

Query Match	90.7%;	Score 201.4;	DB 10;	Length 1063;
Best Local Similarity	95.7%;	Pred. No. 1.9e-24;		
Matches 221; Conservative	0;	Mismatches 1;	Indels 9;	Gaps 1;

[illegible]

RESULT	3
LOCUS	BG107702
DEFINITION	BC107702 1340 bp mRNA
ACCESSION	60227787891 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4365474 5',
VERSION	BG107702
KEYWORDS	BG107702.1 GI:12601548
SOURCE	EST
ORGANISM	human.
	Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukariota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1340)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10014 row: 1 column: 19
 High quality sequence stop: 667.

FEATURES

SOURCE

1. .1340

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4363474"
/clone_lip="NIH_MGC_86"
/tissue_type="osteosarcoma, cell line"
/lab_host="DH10B (phage-resistant)"

```

BASE COUNT	467 a	272 c	450 g	151 t
ORIGIN				

Query Match	90.7%;	Score 201.4;	DB 11;	Length 1340;
Best Local Similarity	95.7%;	Pred. No. 1.9e-24;		
Matches 221;	Conservative	0;	Mismatches 1;	Indels 9;
				Gaps 1;

1 gacatcagcagacacgagagagagagaccacgctcgcgcacgagtcctctcgcacgacgagatcaggc 60
|||||

OY	31	tcggccgagcttcctcgtcaagcaagaatcaagctcaagttcttgtagaagaagagagagagag	90
Db	1191	TCGGCCAAATCACCGTGATGATCTACCACTGCCAGAAGAAGAAGAAGAGAGAGAGAG	1132
OY	91	gaagaaggagagagagagagagagaccacgttgaagaatcagagagagagagagagagag	150
Db	1131	GAGGAAGAAGAGAGGCGAAGTAGAAGAAGTGGAAAGTAGAAGTAGATGCTGAAGTG	1072
OY	151	gagagagagacccgacgcaactctgtgagggagatcaagcagctctgccgaagaagt	206
Db	1071	GAGAACTACTAGATTATTAGTGCAGAGAAGACCTAGATTGGTGTGGCAGAAGAGT	1016

```

RESULT 11
US-08-396-001-13/C
Sequence 13, Application US/08396001
Patent No. 5919618
GENERAL INFORMATION:
APPLICANT: Guarente, Leonard P.
APPLICANT: Austriaco Jr., Nicnor
APPLICANT: Claus, James
APPLICANT: Cole, Francesca
APPLICANT: Kennedy, Brian
TITLE OF INVENTION: Genes Determining Cellular Senescence in
TITLE OF INVENTION: Yeast
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/396,001
FILING DATE: 28-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MT-6408a2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 2150 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 563..1987
US-08-396-001-13

```

Query Match	Similarity	Score	DB 2	Length
Best Local	58.0%	Pred. No. 2e-05		2150
Matches 102	Conservative	0	Mismatches 74	Indels 0
				Gaps 0
QY 31	tcggtccgagtccttcg	tcagcagaa	tcagcagtcctc	taagaagaagagagagag 90
Db 1191	TCGCGCAATCACCCAGATCTACTACCACTGCCAGAAAGACAGAGAGAGAGAGC			1132
QY 91	gaagaggaagaggaagagagagcaccacgttaagaatcagagagagaagagagag 150			

[illegible]

```

RESULT 12
US-09-323-433A-13/c
; Sequence 13: Application US/09323433A
; Patent No. 6218512
; GENERAL INFORMATION:
; APPLICANT: Guarente, Leonard P.
; APPLICANT: Austriaco Jr., Milcator
; APPLICANT: Claus, James J.
; APPLICANT: Cole, Francesca
; APPLICANT: Kennedy, Brian
; TITLE OR INVENTION: GENES DETERMINING CELLULAR SENESENCE IN
; TITLE OR INVENTION: YEAST
; FILE REFERENCE: 0050.1491-003
; CURRENT APPLICATION NUMBER: US/09/323,433A
; CURRENT FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: US 08/396,001
; PRIOR FILING DATE: 1995-02-28
; PRIOR APPLICATION NUMBER: PCT/US94/09351
; PRIOR FILING DATE: 1994-08-15
; PRIOR APPLICATION NUMBER: US 08/107,408
; PRIOR FILING DATE: 1993-08-16
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 2150
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (363)...(1987)
; OTHER INFORMATION: SAG1
; US-09-323-433A-13

```

Query Match	25.9%	Score 57.6	DB 4	Length 2150
Best Local Similarity	58.0%	Pred. No. 2e-05		
Matches 102	Conservative 0	Mismatches 74	Indels 0	Gaps 0

QY	31	tcgagccgagtcctcgcgcacgaacatcagctcaggtctctgaggaagaagaagagag 90
Dc	1191	TCGCGCAAAATCACCCTAGTACTACTACCACTGCCAGAGAAGAAGAAGAAGAGGAGGAG 1132
QY	91	gaagagagagaggaagaagaaggagaccgcctgtaagaatcaagagagaggaagagag 150
Dc	1131	GAGGAAGAGAGAGAGAGAACTAGACAGAGTGTGAAGTGAAGAAGTGAAGTGTGAAGTG 1072
QY	151	gagagagagagccgcgcacactctagagagagcatcagtcagctctgcgcgaagaagt 206
Dc	1071	GAGGAGACTCTAGATTAGTGGCAGAGAAGACACACTAGATTGGTTGGCGAAGAAGT 1016

RESULT 13
US-08-728-323A-1
; Sequence 1, Application US/08728323A
; Patent No. 5948676
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Immediate Early Protein From Kaposi's
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
; TITLE OF INVENTION: Encoding Same And Uses Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:

[illegible]

```

RESULT 10
US-08-861-464-13/c
: Sequence 13, Application US/08861464
: Patent No. 5874210
: GENERAL INFORMATION:
: APPLICANT: Guarente, Leonard P.
: APPLICANT: Austriaco Jr., Micanor
: APPLICANT: Kennedy, Brian
: TITLE OF INVENTION: Genes Determining Cellular Senescence
: TITLE OF INVENTION: in Yeast
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
: STREET: Two Millitia Drive
: CITY: Lexington
: STATE: MA
: COUNTRY: USA
: ZIP: 02173
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/861,464
: FILING DATE: 22-MAY-1997
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/396,001
: FILING DATE: 28-FEB-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US94/09351
: FILING DATE: 15-AUG-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/107,408
: FILING DATE: 16-AUG-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Granahan, Patricia
: REGISTRATION NUMBER: 32,227
: REFERENCE/DOCKET NUMBER: MTT-6408A22
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 781-861-6240
: TELEFAX: 781-861-9540
: INFORMATION FOR SEQ ID NO: 13:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2150 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 563..1987
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 563..1987
:
US-08-861-464-13

```


Qy	33	gcccga	ctctc	gtcag	caagatc	agctc	agcttc	gtctt	cgaagaa	gaagagagagag	92
Dd	662	GGACGAG	AGACGGG	AGAGACG	AGCGGGAGG	AGCGGGAGG	AGCGGGAGG	AGCGGGAGG	AGCGGGAGG	AGCGGGAGG	603
Qy	93	agaagaa	gaagaa	gaagagga	gaaccag	tgaagaa	tcagagag	gaagaa	gaagagagag	152	
Dd	602	GGAGGAG	AGACGAG	CGGGGAG	AGACGAG	CGGGGAG	AGACGAG	CGGGGAG	AGACGAG	CGGGGAG	543
Qy	153	ggagag	gaaccg	gcagca	actctg	aggaag	catcag	agcag	tctgc	ccgaagaa	212
Dd	542	CGGGGAG	AGAGACG	AGCGGGAG	AGACGAG	CGGGGAG	AGACGAG	CGGGGAG	AGACGAG	CGGGGAG	483
Qy	213	ggaaga	218								
Dd	482	GGAAG	477								
<p>RESULT 8 US-08-785-420-1 ; Sequence 1, Application US/08785420 ; Patent No. 6001976 ; GENERAL INFORMATION: ; APPLICANT: MacLennan, David H ; APPLICANT: O'Brien, Peter J ; TITLE OF INVENTION: DIAGNOSIS FOR PORCINE MALIGNANT ; TITLE OF INVENTION: HYPERTHERMIA ; NUMBER OF SEQUENCES: 3</p>											

```

1  RESULT      8
2  US-08-785-420-1
3  ; Sequence 1, Application US/08785420
4  ; Patent No. 6001976
5  ; GENERAL INFORMATION:
6  ; APPLICANT: MacLennan, David H
7  ; APPLICANT: O'Brien, Peter J.
8  ; TITLE OF INVENTION: DIAGNOSIS FOR PORCINE MALIGNANT
9  ; TITLE OF INVENTION: HYPERHEMIA
10 ; NUMBER OF SEQUENCES: 3
11 ; CORRESPONDENCE ADDRESSES:
12 ; ADDRESSEE: Bell, Seltzer, Park & Gibson
13 ; STREET: P.O. Drawer 34009
14 ; CITY: Charlotte,
15 ; STATE: No. 6001976th Carolina 28234
16 ; COUNTRY: U.S.A.
17 ; ZIP: 28234
18 ; COMPUTER READABLE FORM:
19 ; MEDIUM TYPE: Floppy disk
20 ; COMPUTER: IBM PC compatible
21 ; OPERATING SYSTEM: PC-DOS/MS-DOS
22 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
23 ; CURRENT APPLICATION DATA:
24 ; APPLICATION NUMBER: US/08/785,420
25 ; FILING DATE:
26 ; CLASSIFICATION:
27 ; PRIOR APPLICATION DATA:
28 ; APPLICATION NUMBER: 08/249,388
29 ; FILING DATE:
30 ; APPLICATION NUMBER: US 08/030,159
31 ; FILING DATE: 15-MAR-1993
32 ; ATTORNEY/AGENT INFORMATION:
33 ; NAME: Layton, Jr., Samuel G.
34 ; REGISTRATION NUMBER: 22807
35 ; REFERENCE/DOCKET NUMBER: 3477-73
36 ; TELECOMMUNICATION INFORMATION:
37 ; TELEPHONE: 704-377-1561
38 ; TELEFAX: 704-334-2014
39 ; TELEX: 57-5102
40 ; INFORMATION FOR SRO. ID NO.: 1:
41 ; SEQUENCE CHARACTERISTICS:
42 ; LENGTH: 15378 base pairs
43 ; TYPE: nucleic acid
44 ; STRANDEDNESS: single
45 ; TOPOLOGY: linear
46 ; MOLECULE TYPE: DNA (genomic)
47 ; HYPOTHETICAL: NO
48 ; ANTI-SENSE: NO
49 ; IMMEDIATE SOURCE:
50 ; CLONE: Porcine RYR1 Gene
51 ; POSITION IN GENOME:
52 ; UNITS: bp
53 ; US-08-785-420-1

```

Query Match	Score	DB	Length
26.0%;	57.8;	3;	15378;

```
US-09-357-014-8
; Sequence 8, Application US/09357014
; Patent No. 6291645
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Jung, Ratna K. Vadlamudi
; and Jack L. Strominger
; TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
; AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/357,014
; FILING DATE: 19-Jul-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/574,959
; FILING DATE: <UNKNOWN>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragoras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3211 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 439..3157
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-357-014-8

Query Match      27.3% Score 60.6; DB 4; Length 3211;
Best Local Similarity 59.6%; Pred. No. 4.3e-06;
Matches 102; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

OY 48 agcagaatcagctcaggtctgaggaagaagaagaagaagaagaagaagaagaaga 107
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2442 AGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2501
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 108 agggagaccacagtgaagaatcagaggaagaagaagaagaagaagaagaaccgcag 167
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2502 GGAAGAGGATGAAGAGGATATTGTAAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 2561
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 168 caactctgagagcagcatcagagcagctgcggaagaagtaagtgaggaaga 218
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 2562 AGAATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGA 2612

RESULT 5
US-08-574-959A-6
; Sequence 6, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Jung, Ratna K. Vadlamudi
; and Jack L. Strominger
; TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
```

```

; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,959A
; FILING DATE: 19-DEC-95
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragoras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3901 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 439..3847
US-08-574-959A-6

Query Match      27.3% Score 60.6; DB 2; Length 3901;
Best Local Similarity 59.6%; Pred. No. 4.4e-06;
Matches 102; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

OY 48 agcagaatcagctcaggtctgaggaagaagaagaagaagaagaagaagaagaaga 107
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3132 AGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3191
    || ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 108 agggagaccacagtgaagaatcagaggaagaagaagaagaagaagaagaaccgcag 167
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3192 GGAAGAGGATGAAGAGGATATTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3251
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 168 caactctgagagcagcatcagagcagctgcggaagaagtaagtgaggaaga 218
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3252 AGAATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGGA 3302

RESULT 6
US-09-357-014-6
; Sequence 6, Application US/09357014
; Patent No. 6291645
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Jung, Ratna K. Vadlamudi
; and Jack L. Strominger
; TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
; AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

	Query Match	30.4%	Score 67.4;	DB 1;	length 7218;
	Best Local Similarity	0.9%;	Prid. No. 1.3e-07;		
	Matches	2;	Conservative 161;	Mismatches 52;	Indels 0;
QY	7	aagcacagcgaggagaacccacgtcgcgcatcttcgtcaacaatcacgtcagt	66		
DG	1279	RRR	1220		
OY	67	tctgggaagaaagagagagagaagagagagagagagagagagacaccgttaaga	126		
DG	1219	RRR	1160		
OY	127	tcaagagagaaagaaaagagagagagagagacgcgcgacaactctgaagcata	186		
DG	1159	RRR	1100		
OY	187	gagcagctcgcgcgaagaatactaagagaagaat	221		
DG	1099	KKKKKKRRRRRRRRRRRRRRRRRRRRRRRRRRRRAT	1065		

```

1      RESULT      2
2      US-08-781-891-209/c
3      ; Sequence 209 Application US/08781891
4      ; Patent No. 6090620
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Fu, Ying-Hui
7      ; APPLICANT: Yu, Chang-Hu
8      ; APPLICANT: Oshima, Junko
9      ; APPLICANT: Mulligan, John T.
10     ; APPLICANT: Schellenberg, Gerald D.
11     ; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
12     ; NUMBER OF SEQUENCES: 209
13     ; CORRESPONDENCE ADDRESS:
14     ; ADDRESSEE: SEED AND BERRY LLP
15     ; STREET: 6300 Columbia Avenue, 701 Fifth Avenue
16     ; City: Seattle
17     ; STATE: Washington
18     ; COUNTRY: USA
19     ; ZIP: 98104-7092
20     ; COMPUTER READABLE FORM:
21     ; MEDIUM TYPE: floppy disk
22     ; COMPUTER: IBM PC compatible
23     ; OPERATING SYSTEM: PC-DOS/MS-DOS
24     ; SOFTWARE: PatentIn Release #1.0, Version #1.30
25     ; CURRENT APPLICATION DATA:
26     ; APPLICATION NUMBER: US/08/781,891
27     ; FILING DATE: 27-DEC-1996
28     ; CLASSIFICATION: 800
29     ; ATTORNEY/AGENT INFORMATION:
30     ; NAME: No. 6090620lenburg Ph.D., Carol
31     ; REGISTRATION NUMBER: 39,317
32     ; REFERENCE/DOCKET NUMBER: 240052.419
33     ; TELECOMMUNICATION INFORMATION:
34     ; TELEPHONE: (206) 622-4900
35     ; TELEFAX: (206) 682-6031
36     ; INFORMATION FOR SEQ ID NO: 209:
37     ; SEQUENCE CHARACTERISTICS:
38     ; LENGTH: 51259 base pairs
39     ; TYPE: nucleic acid
40     ; STRANDEDNESS: single
41     ; TOPOLOGY: linear
42     ;
43     ; US-08-781-891-209

```

		28.2%	Score 62.6	DB 3	Length 51259.
Query Match Similarity		63.8%	Pred. No. 2.4e-06		
Best Local Similarity					
Matches	95; Conservative	0;	Mismatches 56;	Indels	0; Gaps
OY	70 gaaggaagaaggaggaaggaagaagaagaagaagaagcaccagtgaatca	129			
Dh	226 GAGGAGGAGCAGCAGAGGAGGAGGAGGAAGAAGGAGGAGCGAG	167			

[illegible]

RESULT 3
 US-08-574-959A-8
 Sequence 8 Application US/08574959A
 Patent No. 5962224
 GENERAL INFORMATION:
 APPLICANT: Jaekyoon Shin, Insil Young, Ratna K.
 APPLICANT: and Jack L. Strominger
 TITLE OF INVENTION: P62 POLYPEPTIDES, RELATED PEPTIDES
 TITLE OF INVENTION: AND USES THEREFOR
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHYE & COCKFIELD
 STREET: 60 State Street, Suite 510
 City: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109-1875
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/574,959A
 FILING DATE: 19-DEC-95
 ATTORNEY/AGENT INFORMATION:
 NAME: Mandragouras, Amy E.
 REGISTRATION NUMBER: 36,207
 REFERENCE/DOCKET NUMBER: DFN-008
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)227-7400
 TELEFAX: (617)227-5941
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3211 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 439..3157
 US-08-574-959A-8

[illegible]

AC	AAQ39212;
XX	
DT	21-JUL-1993 (first entry)
XX	
DE	CENP-B cDNA.
XX	
KM	Centromere; autoantigen; anticentromere; ACA; CREST; chromosomal;
KW	auto immune disease; calcinosis; Raynaud's phenomenon; esophageal;
KX	disability; scleroactyl; elanglectasiae; scleroderm; ss.
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FH	CD5
FT	/tag= 1..1785
FT	/tag= a
FT	/tag= b
FT	misc_feature
FT	39..483
FT	/note= "from clone CNP-B3"
FT	484..1334
FT	/tag= c
FT	misc_feature
FT	/note= "from clone CNP-B2"
FT	1335..2643
FT	/tag= d
FT	/note= "from clone CNP-B1"
PX	
PN	USJ196307-A.
XX	
PD	23-MAR-1993.
XX	
PF	29-MAR-1988; 88US-0174854.
XX	
PR	29-MAR-1988; 88US-0174854.
XX	
PA	(UYJO) UNITV JOHNS HOPKINS.
XX	
PI	Cleveland D, Earnshaw W, Sullivan KF;
XX	
DR	WPI; 1993-116791/14.
XX	
P-PSDB:	AAR34936.
PT	Human centromere antigen CENP-B and its cDNA - for detecting
XX	auto-antibodies to CENP-B in auto-immune disease diagnosis
PS	Claim 1; Fig 1; 20pp; English.
XX	
CC	Screening of a human endothelial cell cDNA library in lambda gtl1
CC	with serum GS provided a positive clone, lambda CENP-B1, contg. a
CC	cDNA insert of 1,282 bp. Phage clones contg. longer cDNA inserts
CC	were obt'd. by screening a cDNA library constructed from human
CC	placental mRNA, clones lambda CENP-2-4. The clones were found to
CC	have overlapping sequences and together made up the complete CENP-B
CC	cDNA sequence. The cDNA encodes a centromere polypeptide comprising
CC	an epitope which is reactive with samples contg. anti-centromere
CC	autoantibodies (ACA) and is destroyed by disruption at a site 42
CC	amino acids from the C-terminus of CENP-B or immunologically binds
CC	the monoclonal antibody produced by hybridoma cell line ARCC HB 9667.
CC	The cDNA may be used to transform cells which may be used to produce
CC	recombinant CENP-B which may be used as an assay reagent for detection
CC	of autoantibodies in the diagnosis of autoimmune diseases, esp.
CC	calcinosis/Raynaud's phenomenon/esophageal dysmobility/scleroactyl/
CC	elangiectasiae (CREST) variant of scleroderma.
XX	
SQ	Sequence 2643 BP; 512 A; 801 C; 878 G; 451 T; 1 other;
Query Match	29.1%; Score 64.6; DB 14; Length 2643;
Best Local Similarity	64.2%; Pred. No. 3.5e-06;
Matches 97:	Conservative 0; Mismatches 54; Indels 0; Gaps 0;
DY	70 gaggaagaagaaaggaggaaggaaggaaggaaggaagcaccagtgaaagaatca 129 db 1201 gaggaagaaggaaggaaggaaggaaggaaggaaggaaggtcgaaaggaaggaag 1260

Oy	130	gaggaagaagaagaagaagaagaagaagaacccgcagcacccttcaggagcatcacag	189
Db	1261	gaggaagaagaagaagaagaagaagaagaagggggaagaagaagaattggggagaa	1320
Oy	190	cagtctgccgaagaagtgaagaagaa	220
Db	1321	gaggaagtggaagagagaggtgatgttgate	1351
<hr/>			
RESULT 14			
ID	AAA64660	standard; cDNA; 3717 BP.	
XX	AC	AAA64660;	
XX	DT	02-JAN-2001 (first entry)	
DE	DNA encoding centromere autoantigen B (CENP-B) protein.		
XX	XX	Melanoma; thyroid tumour; MEL3; MEL7; Thy5; Thy6; Thy11; Thy14; Thy15;	
KW	XX	rectal cancer; lung cancer; breast cancer; colon cancer;	
KM	XX	centromere autoantigen B; CENP-B protein; ss.	
OS	Homo sapiens.		
XX	PN	WO200050595-A2.	
XX	PD	31-AUG-2000.	
XX	PF	25-FEB-2000; 2000MO-USO4529.	
XX	PR	25-FEB-1999; 99US-0257417.	
PA	(GOUT/) GOUT I.		
PA	(RODN/) RODNIN N.		
PA	(FILO/) FILOMENKO V.		
PA	(MATS/) MATSUKA G.		
PA	(SCAN/) SCANLAN M.		
PA	(OLDL/) OLD L.		
PA	(BILY/) BILYNSKY B.		
PI	Gout I, Rodnin N, Filonenko V, Matsuka G, Scanlan M, Old L;		
PI	Bilynsky B;		
DR	WPI: 2000-572092/53.		
XX			
PT	Novel isolated nucleic acid molecules for diagnosing and treating		
XX	melanoma, thyroid tumors, rectal, lung, breast and colon cancers -		
PS	Claim 17: Page 86-87; 94pp; English.		
CC	The specification describes polynucleotides which are associated with		
CC	melanoma and thyroid tumors. Specifically, these are MEL3, MEL7, Thy5,		
CC	Thy6, Thy11, Thy14, and Thy15. The polynucleotides are useful for		
CC	diagnosing and treating a patient with melanoma, thyroid tumour,		
CC	rectal cancer, lung cancer, breast cancer or colon cancer. The		
CC	present sequence represents a polynucleotide of the invention.		
SQ	Sequence 3717 BP; 705 A; 1133 C; 1214 G; 665 T; 0 other:		
<hr/>			
Query Match 29.1%; Score 64.6; DB 21; Length 3717;			
Best Local Similarity 64.2%; Pred. No. 3.6e-06;			
Matches 97; Conservative 0; Mismatches 54; Indels 0; Gaps 0;			
Oy	70	gaggaagaagaagaagaagaagaagaagaagcgagcagcactctggaggatcatcag	129
Db	2151	gaggaagaagaagaagaagaagaagaagaagggtgaagaagaagaag	2210
Oy	130	gaggaagaagaagaagaagaagaagaagaccgcagcagcactctggaggatcatcag	189
Db	2211	gaggaagaagaagaagaagaagaagaagaccgcagcagcactctggaggatcatcag	2270

KM	spondyloepiphyseal dysplasia; achondroplasia; human; ds.
XX	
XX	Homo sapiens.
PN	NC_00950284-A2.
PN	
XX	07-OCT-1999.
PD	
XX	
PF	26-MAR-1999; 99WO-EP02055.
XX	
PR	27-MAR-1998; 98DE-1013799.
XX	(ROSE/) ROSENTHAL A.
PA	
XX	Rosenthal A, Rump A, Hess J, Aigner T, Wirth T;
PI	
XX	WPI; 1999-601320/51.
DR	
XX	
PT	Nucleic acids encoding proteins which influence bone development,
XX	useful for treating and studying bone disorders -
PS	
XX	Example 3; Page 272-300; 391pp; German.
CC	This invention describes novel nucleic acids (I; designated LOBO (long
CC	bones)) encoding proteins influencing bone development in mammals. The
CC	proteins of the invention reduce and/or inactivate bone extension (i.e.,
CC	development), with exception of the skull and have osteopathic activity.
CC	The nucleic acid molecules, proteins and antibodies can be used in
CC	diagnostic or pharmaceutical compounds e.g. for gene therapy. The methods
CC	and nucleic acid molecules, etc. are useful for production of transgenic
CC	animals, especially a transgenic mouse for the study of diseases
CC	associated with bone development, e.g. spondyloepiphyseal dysplasia and
CC	achondroplasia. This sequence encodes a human LOBO protein described
CC	in the method of the invention.
XX	
SQ	Sequence 49999 BP; 12607 A; 13453 C; 13304 G; 10635 T; 0 other;
XX	
Query Match	29.5%; Score 65.4; DB 20; Length 49999;
Best Local Similarity	68.9%; Pred. No. 3.5e-06;
Matches 104; Conservative	0; Mismatches 46; Indels 1; Gaps 1.
OY	70 gaggaagaagaggaggaagaagagggaggaaggaagcaccgtgaagatca 129
DB	49833 gaggaagaagaggaggaagaggagggaggaaggaagag-gaggaagaagaagaag 49891
OY	130 gaggaagaaggaagaaggaaggaagggaggaacccgaaactctaggagcatcaag 189
DB	49892 gaggaagaaggaagaaggaaggaaggaagaagaagaagaagaaggaagaagga 49951
OY	190 cagtcgcggaagaagttaagtgaagaaga 220
DB	49952 gaggaagaaggaagaagaagaagaagaaga 49982
RESULT 12	
AAQ10613	
ID	AAQ10613 standard; DNA; 15672 BP.
XX	
AC	AAQ10613;
XX	
DT	26-APR-1991 (first entry)
XX	
DE	Rianodin receptor gene.
XX	
KM	Rianodin receptor gene; calcium release modulator; tranquilliser; antagonist; ss.
XX	
OS	Oryctolagus cuniculus.
XX	
Key	Location/Qualifiers
FH	593..15553
FT	/tag= a

RESULT	10
AA223896	
ID	AA223896 standard; DNA; 4999 BP.
XX	
AC	AA223896;
XX	

[illegible]

RESULT	11
AAZ23901	
ID	AAZ23901 standard; DNA; 49999 BP.
XX	
AC	AAZ23901;
XX	
DT	25-JAN-2000 (first entry)
XX	
DE	Human LOBO homologue genomic DNA fragment 3.
XX	
LOBO;	long bones; bone development; bone extension; skull; osteopathic;
KW	diagnostic; pharmaceutical; gene therapy; transgenic animal; disease;

```
FT /note= "CDS does not include start codon"
FT /partial
FT mutation (931..932, AGAG)
FT /*tag= b
FT /note= "This deletion results in the change of
FT amino acids and ends in a premature stop codon"
XX
XX WO200138578-A1.
XX
XX 31-MAY-2001.
XX
XX 21-NOV-2000; 2000WO-US31940.
XX
XX 24-NOV-1999; 99US-0167365.
XX
XX (CORR ) CORNELL RES FOUND INC.
XX
XX Aguirre GD, Acland GM, Zhang Q, Ray K, Zeiss CJ;
XX
XX WPI: 2001-367707/38.
XX
XX P-PSDB; AAE02399.
XX
XX Identifying dogs with or carrying X-linked progressive retinal atrophy
XX by detecting retinitis pigmentosa GTPase regulator gene mutation,
XX useful when breeding Husky, Samoyed and Miniature Schnauzer -
XX
XX Claim 75; Page 34-35; 88pp; English.
XX
XX The invention relates to a method for identifying dogs which are
XX genetically normal, are carriers of, or are affected with X-linked
XX progressive retinal atrophy (XLPRA), by testing a biological sample with
XX genetic markers that co-segregate with a XLPRA gene locus. The invention
XX also relates to canine retinitis pigmentosa GTPase regulator (RPRG). RPRG
XX mutants and their corresponding nucleic acid molecules. The mutated RPRG
XX genes are responsible for the XLPRA in dogs. Methods are used to select
XX dogs for breeding so that dogs carrying the mutated locus are eliminated
XX from the breeding stock. The method particularly applies to Siberian
XX Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
XX the disease is X-linked. XLPRA1 type is identified in Miniature Schnauzers,
XX and Samoyeds, while XLPRA2 type is identified in Siberian Huskies.
XX The present cDNA sequence is the mutant open reading frame (ORF) 15
XX encoding Canine retinitis pigmentosa GTPase regulator (RPRG) mutant
XX found in XLPRA2-affected dogs. This mutant is obtained by deleting
XX 'GA' nucleotide bases from position 932 to 933 of the wild-type
XX canine RPRG cDNA.
XX
XX Sequence 2803 BP; 1028 A; 287 C; 939 G; 549 T; 0 other:
SQ
Query Match 30.6%; Score 68; DB 22; Length 2803;
Best Local Similarity 62.2%; Pred. No. 5.9e-07;
Matches 107; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
```

```
XX
XX Canine retinitis pigmentosa GTPase regulator (RPRG) cDNA.
XX
XX Dog: X-linked progressive retinal atrophy; XLPRA; genetic marker;
XX retinitis pigmentosa GTPase regulator; RPRG; Siberian Husky; Samoyed;
XX Miniature Schnauzer; ss.
XX
XX Canis familiaris.
XX
XX Key Location/Qualifiers
XX CDS 1..1569
XX
XX /*tag= a
XX /product= "Canine retinitis pigmentosa GTPase regulator"
XX /note= "CDS does not include start codon"
XX
XX WO200138578-A1.
XX
XX 31-MAY-2001.
XX
XX 21-NOV-2000; 2000WO-US31940.
XX
XX 24-NOV-1999; 99US-0167365.
XX
XX (CORR ) CORNELL RES FOUND INC.
XX
XX Aguirre GD, Acland GM, Zhang Q, Ray K, Zeiss CJ;
XX
XX WPI: 2001-367707/38.
XX
XX P-PSDB; AAE02399.
XX
XX Identifying dogs with or carrying X-linked progressive retinal atrophy
XX by detecting retinitis pigmentosa GTPase regulator gene mutation,
XX useful when breeding Husky, Samoyed and Miniature Schnauzer -
XX
XX Claim 71; Page 30-31; 88pp; English.
XX
XX The invention relates to a method for identifying dogs which are
XX genetically normal, are carriers of, or are affected with X-linked
XX progressive retinal atrophy (XLPRA), by testing a biological sample with
XX genetic markers that co-segregate with a XLPRA gene locus. The invention
XX also relates to canine retinitis pigmentosa GTPase regulator (RPRG). RPRG
XX mutants and their corresponding nucleic acid molecules. The mutated RPRG
XX genes are responsible for the XLPRA in dogs. Methods are used to select
XX dogs for breeding so that dogs carrying the mutated locus are eliminated
XX from the breeding stock. The method particularly applies to Siberian
XX Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
XX the disease is X-linked. XLPRA1 type is identified in Miniature Schnauzers,
XX and Samoyeds, while XLPRA2 type is identified in Siberian Huskies.
XX The present cDNA sequence is the normal open reading frame (ORF) 15
XX encoding Canine retinitis pigmentosa GTPase regulator (RPRG) protein
XX found in normal dogs.
XX
XX Sequence 2805 BP; 1029 A; 287 C; 940 G; 549 T; 0 other:
SQ
Query Match 30.6%; Score 68; DB 22; Length 2805;
Best Local Similarity 62.2%; Pred. No. 5.9e-07;
Matches 107; Conservative 0; Mismatches 65; Indels 0; Gaps 0;
```

PR 24-NOV-1999; 99US-0167365.

```

FT      /tag= a
FT      /product= "Canine retinitis pigmentosa GTPase
FT      regulator mutant"

```

```
XX
PI Cornells S, Beyaert R;
XX
DR WPI; 2000-499331/44.
XX
PT Nucleic acids encoding internal ribosome entry sequences useful for
XX directing protein expression in gene therapy procedures -
XX
PS Claim 10; Page 34; 57pp; English.
XX
CC The PITSLRE protein kinase gene can be translated to give two isoforms,
XX p110 and p58. Transcription of p58 mRNA was found to start at an
XX internal ribosome entry site (IRES). The IRES element was found to be
XX cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX cycle. The present sequence contains the PITSLRE protein kinase
XX IRES and upstream nucleotides. The IRES sequence and fragments of it
XX may be used to induce cell cycle dependent initiation of translation in
XX eukaryotic cells. Vectors containing the IRES may be used for the
XX preparation of compositions for the treatment of and/or prevention of
XX a disease by gene therapy. Such diseases may be cancer and restenosis.
XX
SQ Sequence 660 BP; 210 A; 128 C; 261 G; 61 T; 0 other;

Query Match      100.0%; Score 222; DB 21; Length 660;
Best Local Similarity 100.0%; Pred. No. 4.3e-42;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagacagaggaagaccagctcgcgcgagctcctcgtcagacaatcagc 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 439 gactcagcgacagcagcagaggaagaccagctcgcgcgagctcctcgtcagacaatcagc 498

QY 61 tcaggttcctgaggaagaagagaggaagaggaagaggaagaggaagaggaagcaccagt 120
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 499 tcaggttcctgaggaagaagaggaagaggaagaggaagaggaagaggaagaggaagcaccagt 558

QY 121 gaagaatcagagaggaagaggaagaggaagaggaagaggaagaggaagaggaag 180
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 559 gaagaatcagagaggaagaggaagaggaagaggaagaggaagaggaagaggaag 618

QY 181 gcatcagagcagctcctcgcgaagaagtaagtgaaggaagaatg 222
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 619 gcatcagagcagctcctcgcgaagaagtaagtgaaggaagaatg 660

RESULT 4
AAAT3711
ID AAA73711 standard; DNA; 2471 BP.
XX
AC AAA73711;
XX
DT 07-DEC-2000 (first entry)
XX
DE Human PITSLRE protein kinase isoform alpha2-2.
XX
KW IRES; internal ribosome entry site; PITSLRE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_signal 907..1128
FT /*tag= a
FT /label= IRES
FT /note= "Internal ribosome entry site"

MO200044896-A1.
XX
PN 03-AUG-2000.
XX
PD 26-JAN-2000; 2000MO-EP00643.
XX
PR 26-JAN-1999; 99EP-0200216.
XX
```

```
XX
PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI Cornells S, Beyaert R;
XX
DR WPI; 2000-499331/44.
XX
PT Nucleic acids encoding internal ribosome entry sequences useful for
XX directing protein expression in gene therapy procedures -
XX
PS Disclosure; Page 32-33; 57pp; English.
XX
CC The PITSLRE protein kinase gene can be translated to give two isoforms,
XX p110 and p58. Transcription of p58 mRNA was found to start at an
XX internal ribosome entry site (IRES). The IRES element was found to be
XX cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX cycle. The present sequence contains the human PITSLRE protein kinase isoform
XX alpha2-2. This sequence contains the IRES. The IRES sequence and
XX fragments of it may be used to induce cell cycle dependent initiation
XX of translation in eukaryotic cells. Vectors containing the IRES may be
XX used for the preparation of compositions for the treatment of and/or
XX prevention of a disease by gene therapy. Such diseases may be cancer
XX and restenosis.
XX
SQ Sequence 2471 BP; 697 A; 601 C; 809 G; 364 T; 0 other;

Query Match      100.0%; Score 222; DB 21; Length 2471;
Best Local Similarity 100.0%; Pred. No. 5.2e-42;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagacagcagaggaagaccagctcgcgcgagctcctcgtcagacaatcagc 60
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 907 gactcagcgacagcagcagaggaagaccagctcgcgcgagctcctcgtcagacaatcagc 966

QY 61 tcaggttcctgaggaagaagaggaagaggaagaggaagaggaagaggaagaggaagcaccagt 120
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 967 tcaggttcctgaggaagaagaggaagaggaagaggaagaggaagaggaagaggaagcaccagt 1026

QY 121 gaagaatcagaggaagaagaagaggaagaggaagaggaagaggaagaggaagaggaag 180
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1027 gaagaatcagaggaagaagaagaggaagaggaagaggaagaggaagaggaagaggaag 1086

QY 181 gcatcagagcagctcctcgcgaagaagtaagtgaaggaagaatg 222
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
DB 1087 gcatcagagcagctcctcgcgaagaagtaagtgaaggaagaatg 1128

RESULT 5
AAAT3715
ID AAA73715 standard; DNA; 87 BP.
XX
AC AAA73715;
XX
DT 07-DEC-2000 (first entry)
XX
DE Functional fragment of human PITSLRE protein kinase gene IRES.
XX
KW IRES; internal ribosome entry site; PITSLRE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_signal 907..1128
FT /*tag= a
FT /label= IRES
FT /note= "Internal ribosome entry site"

MO200044896-A1.
XX
PN 03-AUG-2000.
XX
PD 26-JAN-2000; 2000MO-EP00643.
XX
PR 26-JAN-1999; 99EP-0200216.
XX
PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
```

XX Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -
XX
PS Claim 4; Page 32; 57pp; English.
XX
CC The PITSURE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence is the PITSURE protein kinase IRES. The
CC IRES sequence and fragments of it may be used to induce cell cycle
CC dependent initiation of translation in eukaryotic cells. Vectors
CC containing the IRES may be used for the preparation of compositions for
CC the treatment of and/or prevention of a disease by gene therapy. Such
CC diseases may be cancer and restenosis.
XX
SQ Sequence 222 BP; 77 A; 36 C; 89 G; 20 T; 0 other;

Query Match 100.0%; Score 222; DB 21; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.7e-42;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcgacgcgagaggaagaccagctcgccgagctcctcgtagcagaatcagc 60
|||||
Db 1 gacatcagcgacgcgagaggaagaccagctcgccgagctcctcgtagcagaatcagc 60
QY 61 tcaggtctcgaagaagaagagaggaagaggaagaggaagaggaagaggaagaggaag 120
|||||
Db 61 tcaggtctcgaagaagaagagaggaagaggaagaggaagaggaagaggaagaggaag 120
QY 121 gaagaaatcagagaggaagaagaagaagaagaagaagaccgcagcaactctgagag 180
|||||
Db 121 gaagaaatcagagaggaagaagaagaagaagaagaagaccgcagcaactctgagag 180
QY 181 gcatcagagcagctcgcgaagaagtaagtgaagaagaatg 222
|||||
Db 181 gcatcagagcagctcgcgaagaagtaagtgaagaagaatg 222

RESULT 2
AA73710
ID AA73710 standard; RNA; 222 BP.

XX AC AA73710;

XX DT 07-DEC-2000 (first entry)

DE RNA of human PITSURE protein kinase gene internal ribosome entry site.

KW IRES; internal ribosome entry site; PITSURE; human; gene therapy;
KW cancer; restenosis; p58; p110; protein kinase; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
FT misc_signal 1..222
FT /*tag= a
FT /label= IRES
FT /note= "Internal ribosome entry site"

PN MO200044896-A1.

PD 03-AUG-2000.

PF 26-JAN-2000; 2000WO-EP00643.

PR 26-JAN-1999; 99EP-0200216.

PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX Cornells S, Beyaert R;
PI

XX WPI; 2000-499331/44.
XX
XX Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -
XX
PS Disclosure; Page 32; 57pp; English.
XX

CC The PITSURE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence is the corresponding RNA of the PITSURE
CC IRES described in AA73710. The IRES sequence and fragments of it may be
CC used to induce cell cycle dependent initiation of translation in
CC eukaryotic cells. Vectors containing the IRES may be used for the
CC preparation of compositions for the treatment of and/or prevention of
CC a disease by gene therapy. Such diseases may be cancer and restenosis.
XX
SQ Sequence 222 BP; 77 A; 36 C; 89 G; 20 U; 0 other;

Query Match 100.0%; Score 222; DB 21; Length 222;
Best Local Similarity 91.0%; Pred. No. 3.7e-42;
Matches 202; Conservative 20; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcgacgcgagaggaagaccagctcgccgagctcctcgtagcagaatcagc 60
|||||
Db 1 gacatcagcgacgcgagaggaagaccagctcgccgagctcctcgtagcagaatcagc 60
QY 61 tcaggtctcgaagaagaagaggaagaggaagaggaagaggaagaggaagaggaag 120
|||||
Db 61 ucaaguuucugaggaagaggaagaggaagaggaagaggaagaggaagaggaagag 120
QY 121 gaagaaatcagagaggaagaagaagaagaagaagaagaccgcagcaactctgagag 180
|||||
Db 121 gaagaaatcagagaggaagaagaagaagaagaagaagaccgcagcaactctgagag 180
QY 181 gcatcagagcagctcgcgaagaagtaagtgaagaagaatg 222
|||||
Db 181 gcatcagagcagctcgcgaagaagtaagtgaagaagaatg 222

RESULT 3
AA73714
ID AA73714 standard; DNA; 660 BP.

XX AC AA73714;

XX DT 07-DEC-2000 (first entry)

DE Human PITSURE protein kinase gene IRES and upstream sequence.

KW IRES; internal ribosome entry site; PITSURE; human; gene therapy;
KW cancer; restenosis; p58; p110; protein kinase; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
FT misc_signal 439..660
FT /*tag= a
FT /label= IRES
FT /note= "Internal ribosome entry site"

PN MO200044896-A1.

PD 03-AUG-2000.

PF 26-JAN-2000; 2000WO-EP00643.

PR 26-JAN-1999; 99EP-0200216.

PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 17:07:19 ; Search time 355.05 Seconds
(without alignments)
536.055 Million cell updates/sec

Title: US-09-915-060-1

Perfect score: 222

Sequence: 1 gacatcagcagcagcagag.....aagtaagtgcgaagaatg 222

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N.Geneseq_1101: *
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	222	100.0	222	21	AAA73709 Human PITSURE prot
2	222	100.0	222	21	AAA73710 RNA of human PITSURE
3	222	100.0	660	21	AAA73714 Human PITSURE prot
4	222	100.0	2471	21	AAA73711 Human PITSURE prot
5	87	39.2	87	21	AAA73715 Functional fragmen
6	68	30.6	2800	22	AA07131 Canine retinitis p
7	68	30.6	2803	22	AA07132 Canine retinitis p
8	68	30.6	2805	22	AA07130 Canine retinitis p
9	66.8	30.1	49999	20	AA23881 Murine LOBO genom
10	66.8	30.1	49999	20	AA23896 Murine LOBO genom
11	65.4	29.5	49999	20	AA23901 Human LOBO homolog

12	65	29.3	15672	12	AA01613
13	64.6	29.1	2643	14	AA039212
14	64.6	29.1	3717	21	AAA64660
15	64.2	28.9	470	22	AA115222
16	64.2	28.9	470	22	AA133875
17	64.2	28.9	470	22	AA102430
18	62.6	28.2	575	22	AA123315
19	62.6	28.2	575	22	AA148636
20	62.6	28.2	575	22	AA108956
21	62.6	28.2	1969	22	AA114105
22	62.6	28.2	1969	22	AA135486
23	62.6	28.2	1969	22	AA103958
24	62.6	28.2	51259	18	AA03007
25	62	27.9	349	22	AA14995
26	62	27.9	349	22	AA16341
27	62	27.9	349	22	AA104761
28	61.6	27.7	3399	17	AA05868
29	60.6	27.3	3211	18	AA089346
30	60.6	27.3	3901	18	AA089345
31	60.4	27.2	209	22	AA124110
32	60.4	27.2	209	22	AA149402
33	60.4	27.2	209	22	AA109683
34	59.8	26.9	299	22	AA124183
35	59.8	26.9	299	22	AA149470
36	59.8	26.9	299	22	AA109746
37	58.4	26.3	438	22	AA136482
38	58.4	26.3	438	22	AA104897
39	58.4	26.3	117213	19	AA062176
40	58.2	26.2	276	22	AA123164
41	58.2	26.2	2784	18	AA065102
42	58	26.1	475	22	AA11457
43	58	26.1	475	22	AA13728
44	58	26.1	475	22	AA101373
45	58	26.1	512	22	AA106370

ALIGNMENTS

RESULT 1	AAA73709	standard; DNA; 222 BP.
ID	AAA73709	
AC	AAA73709;	
XX		
DT	07-DEC-2000	(first entry)
XX		
DE	Human PITSURE protein kinase gene internal ribosome entry site.	
XX		
KW	IRES; internal ribosome entry site; PITSURE; human; gene therapy;	
KW	cancer; retinosis; p58; p110; protein kinase; ds.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	misc_signal	1..222
FT		/*tag= a
FT		/label= IRES
FT		/note= "internal ribosome entry site"
XX		
PN	WO200044896-A1.	
XX		
PD	03-AUG-2000.	
XX		
PF	26-JAN-2000; 2000WO-EP00643.	
XX		
PR	26-JAN-1999; 99EP-0200216.	
XX		
PA	(VIAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.	
XX		
PI	Cornelis S, Beyaert R;	
XX		
DR	WPI; 2000-499331/44.	

Rianodin receptor
CENP-B CDNA. Homo
DNA encoding cent
Probe #2453 for ge
Probe #2511 used t
Probe #2421 used t
Probe #13248 for g
Probe #17332 used
Probe #8947 used t
Probe #4038 for ge
Probe #4172 used t
Probe #3949 used t
Partial mouse WRN
Probe #4928 for ge
Probe #5027 used t
Probe #4752 used t
Chicken leucocytos
Human p160 CDNA 16
Human p160 CDNA 16
Probe #14043 for g
Probe #18088 used
Probe #674 used t
Probe #14116 for g
Probe #18156 used
Probe #9737 used t
Probe #5168 used t
Probe #4888 used t
HSV-2 strain SB5 C
Probe #13097 for g
Luciferase gene un
Probe #1390 for ge
Probe #1414 used t
Probe #1364 used t
Probe #6361 used t

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144806.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159350.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 66.7%; Score 20; DB 21; Length 831;
Best Local Similarity 82.1%; Pred. No. 66;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2 acatcagcgacagcgagaggaagaccag 29
||||| | ||||| ||| |||||
Db 53 acatcacacacagcaagaacaagaccag 80

Search completed: December 4, 2001, 19:03:36
Job time: 6977 sec


```

XX Erythropoietin; Diamond Blackfan anaemia; polycythemia vera; ss.
KW
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FH sig_peptide 28..99 /tag= a
FT mat_peptide 100..1548 /tag= b
FT polyA_site 1712..1740 /tag= c
FT
XX
XX WO9008822-A.
XX
XX
XX PD 09-AUG-1990.
XX
XX PE 01-FEB-1990; 90WO-US000635.
XX
XX PR 03-FEB-1989; 89US-0306503.
XX
XX PA (GENE-) GENETICS INST INC.
XX PA (WHIT-) WHITEHEAD INST.
XX
XX D'andrea A, Wong G;
XX
XX WP1; 1990-260931/34.
XX DR P-PSDB; AAR06511.
XX
XX Erythropoietin receptor and gene - used for developing reagents
XX and systems to control and study erythropoiesis.
XX
XX
XX Claim 7; Fig 1; 53pp; English.
XX
XX The sequence was obt'd. from a clone isolated from a cDNA library
XX prep'd. from uninduced murine erythroleukemia (MEL) cells. The
XX sequence encodes a type I transmembrane protein with binding
XX affinity for EPO. The gene and recombinant EPO receptor produced
XX on expression of the DNA are used to develop reagents and systems
XX to control and study erythropoiesis. It is believed that the EPO
XX receptor is dysfunctional in individuals with Diamond Blackfan
XX anemia, and may be hyperactive in polycythemia vera.
XX See also AAQ05748 (human EPO receptor clone).
XX
XX
XX Sequence 1740 BP; 352 A; 527 C; 489 G; 372 T; 0 other;
XX
XX
XX Query Match 68.7%; Score 20.6; DB 11; Length 1740;
XX Best Local Similarity 85.2%; Pred. No. 42;
XX Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
XX 3 catcagcgacgagcgaggaagaccag 29
XX | | | | | | | | | | | | | | | | | | | |
XX | | | | | | | | | | | | | | | | | | | |
XX Db 825 CGTCAGCAACAGCGAGTGAGCGACGACG 799
XX
XX
XX RESULT 12
XX ID AAQ053994/c CDNA; 1741 BP.
XX
XX AAQ053994;
XX
XX 24-JUN-1994 (first entry)
XX
XX MEL EPO receptor cDNA.
XX
XX Erythropoietin receptor; recombinant; murine; anaemia; ss.
XX
XX Mus musculus.
XX
XX Key Location/Qualifiers
XX CDS 28..1551
XX /tag= a
XX

```

FT	sig_peptide	28..100
FT		/*tag= b
XX		
PN	US5278065-A.	
PD	11-JAN-1994.	
XX		
PF	03-FEB-1989;	89US-0306503.
XX		
PR	03-FEB-1989;	89US-0306503.
PR	25-MAR-1991;	91US-0678877.
XX		
PA	(CHIL-) CHILDRENS MEDICAL CENT.	
PA	(GEMY) GENETICS INST INC.	
PA	(MHED) WHITEHEAD INST BIOMEDICAL RES.	
PI	D'andrea A, Jones SS, Wong GG;	
XX		
DR	WPI; 1994-025409/03.	
DR	P-PsDB; AAR47517.	
XX		
PT	Recombinant DNA encoding erythropoietin receptor - used to develop prods. for study, treatment or diagnosis of disorders in which receptor is dysfunctional	
PS	Claim 1; Fig 2; 24pp; English.	
XX		
CC	Mouse erythroleukaemia (MEL) cells were used to construct a CDNA library. The CDNA was used to transfect COS-1 cells and these were screened for radiiodinated erythropoietin (Epo) binding to isolate CC CDNA encoding the Epo receptor. The CDNA may be used to isolate CC Epo receptor from other sources and to study, treat or diagnose disorders in which the Epo receptor is dysfunctional.	
CC	See also AAQ53995.	
XX		
SQ	Sequence 1741 BP; 353 A; 528 C; 488 G; 372 T; 0 other:	
	Query Match	68.7%; Score 20.6; DB I5; Length 1741;
	Best Local Similarity	85.2%; Pred. No. 42;
	Matches 23; Conservative	0; Mismatches 4; Indels 0; Gaps 0;
QY	3 catcagcgacgacgaggaagaccag 29	
	I I I I I I I I I I I I I I I I I	
DB	825 CGTCAGCAACAGCAGATGAGGCCAG 799	
	RESULT 13	
	AAQ44854/C	
ID	AAQ44854 standard; CDNA, 1741 BP.	
XX		
AC	AAQ44854;	
XX		
DT	19-OCT-1994 (first entry)	
XX		
DE	Mouse soluble EPO receptor protein gene.	
XX		
KW	Murine; soluble; erythropoietin; EPO; receptor protein; sEPO-R; drug; antigen; diagnostic agent; biochemical reagent; ss.	
XX		
OS	Mus musculus.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	28..1548
FT		/*tag= a
FT	/product= EPO-R	
XX		
PN	JP06038787-A.	
XX		
PD	15-FEB-1994.	
XX		
PT	04-MAR-1992;	92JP-0082865.
XX		

CC diseases may be cancer and restenosis. The present sequence is the
CC antisense primer di-4 used in creation of a PISURE distronic vector.
CC This was used in the process of identifying the IRBS.

XX Sequence 35 BP; 3 A; 12 C; 8 G; 12 T; 0 other;

Query Match 93.3%; Score 28; DB 21; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagcagc 28
|||
Db 35 GACATCAGCGACGACGAGAGAGACCA 8

RESULT 9
AA18242
ID AA18242 standard; DNA; 522 BP.

AC AA18242;

DT 12-OCT-2001 (first entry)

DE Probe #8175 for gene expression analysis in human cervical cell sample.

KW Probe: human; microarray; gene expression; cervical epithelial cell;

KW cervical cancer; ss.

OS Homo sapiens.

PN W0200157278-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00670.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PS (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPL; 2001-488901/53.

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -

PS Claim 25; SEQ ID NO 8175; 487bp; English.

XX The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human Hela cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 522 BP; 107 A; 145 C; 161 G; 109 T; 0 other;

Query Match 68.7%; Score 20.6; DB 22; Length 522;
Best Local Similarity 85.2%; Pred. No. 37;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagcagc 27
|||
Db 312 gacatcagcagcagcagcagcagcagcagc 338

RESULT 10

ID AA143284 standard; DNA; 522 BP.

AC AA143284;

DT 17-OCT-2001 (first entry)

DE Probe #11970 used to measure gene expression in human placenta sample.

KW Probe; microarray; human; placenta; antenatal diagnosis;

KW genetic disorder; ss.

OS Homo sapiens.

PN W0200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PS (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPL; 2001-488897/53.

PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -

PS Claim 25; SEQ ID NO 11970; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs).

CC The present sequence is one such probe. The probes are useful for

CC producing a microarray for predicting, measuring and displaying gene

CC expression in samples derived from human placenta. The probes are useful

CC for antenatal diagnosis of human genetic disorders.

XX Sequence 522 BP; 107 A; 145 C; 161 G; 109 T; 0 other;

Query Match 68.7%; Score 20.6; DB 22; Length 522;
Best Local Similarity 85.2%; Pred. No. 37;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagcagc 27
|||
Db 312 gacatcagcagcagcagcagcagcagcagc 338

RESULT 11

ID AA05747/C standard; cDNA; 1740 BP.

AC AA05747;

DT 04-JAN-1991 (first entry)

DE EPO receptor clone 190.

```
PF 26-JAN-2000; 2000WO-EP00643.
XX
XX 26-JAN-1999; 99EP-0200216.
XX
XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
XX Cornelis S, Beyaert R;
XX
XX WPI; 2000-499331/44.
XX
XX Nucleic acids encoding internal ribosome entry sequences useful for
XX directing protein expression in gene therapy procedures -
XX
XX Clatm 10; Page 34; 57pp; English.
XX
XX The PITSLRE protein kinase gene can be translated to give two isoforms,
XX p110 and p58. Transcription of p58 mRNA was found to start at an
XX internal ribosome entry site (IRES). The IRES element was found to be
XX cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX cycle. The present sequence contains the PITSLRE protein kinase
XX IRES and upstream nucleotides. The IRES sequence and fragments of it
XX may be used to induce cell cycle dependent initiation of translation in
XX eukaryotic cells. Vectors containing the IRES may be used for the
XX preparation of compositions for the treatment of and/or prevention of
XX a disease by gene therapy. Such diseases may be cancer and restenosis.
XX
XX Sequence 660 BP; 210 A; 128 C; 261 G; 61 T; 0 other;

Query Match 100.0%; Score 30; DB 21; Length 660;
Best Local Similarity 100.0%; Pred. No. 0.0095;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcgacagcgaggaagaccagc 30
   |||
Db 439 gacatcagcgacagcgaggaagaccagc 468

RESULT 7
AAA73711
ID AAA73711 standard; DNA; 2471 BP.
XX
XX AAA73711;
XX
XX 07-DEC-2000 (first entry)
XX
XX Human PITSLRE protein kinase isoform alpha2-2.
XX
XX IRES; internal ribosome entry site; PITSLRE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX misc_signal 907..1128
XX /tag= a
XX /label= IRES
XX /note= "internal ribosome entry site"

WO200044896-A1.
XX
XX 03-AUG-2000.
XX
XX 26-JAN-2000; 2000WO-EP00643.
XX
XX 26-JAN-1999; 99EP-0200216.
XX
XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
XX Cornelis S, Beyaert R;
XX
XX WPI; 2000-499331/44.
XX
```

```
PT Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -
XX
XX Disclosure; Page 32-33; 57pp; English.
XX
XX The PITSLRE protein kinase gene can be translated to give two isoforms,
XX p110 and p58. Transcription of p58 mRNA was found to start at an
XX internal ribosome entry site (IRES). The IRES element was found to be
XX cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX cycle. The present sequence is the human PITSLRE protein kinase isoform
XX alpha2-2. This sequence contains the IRES. The IRES sequence and
XX fragments of it may be used to induce cell cycle dependent initiation
XX of translation in eukaryotic cells. Vectors containing the IRES may be
XX used for the preparation of compositions for the treatment of and/or
XX prevention of a disease by gene therapy. Such diseases may be cancer
XX and restenosis.
XX
XX Sequence 2471 BP; 697 A; 601 C; 809 G; 364 T; 0 other;

Query Match 100.0%; Score 30; DB 21; Length 2471;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcgacagcgaggaagaccagc 30
   |||
Db 907 gacatcagcgacagcgaggaagaccagc 936

RESULT 8
AAA73726/C
ID AAA73726 standard; DNA; 35 BP.
XX
XX AAA73726;
XX
XX 07-DEC-2000 (first entry)
XX
XX Antisense primer di-4 used in creation of PITSLRE dicistronic vector.
XX
XX IRES; internal ribosome entry site; PITSLRE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; antisense; PCR primer;
XX ss.
XX
XX Homo sapiens.
XX
XX Synthetic.
XX
XX WO200044896-A1.
XX
XX 03-AUG-2000.
XX
XX 26-JAN-2000; 2000WO-EP00643.
XX
XX 26-JAN-1999; 99EP-0200216.
XX
XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
XX Cornelis S, Beyaert R;
XX
XX WPI; 2000-499331/44.
XX
XX Nucleic acids encoding internal ribosome entry sequences useful for
XX directing protein expression in gene therapy procedures -
XX
XX Examples; Page 25; 57pp; English.
XX
XX The PITSLRE protein kinase gene can be translated to give two isoforms,
XX p110 and p58. Transcription of p58 mRNA was found to start at an
XX internal ribosome entry site (IRES). The IRES element was found to be
XX cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX cycle. The IRES sequence and fragments of it may be used to induce cell
XX cycle dependent initiation of translation in eukaryotic cells. Vectors
XX containing the IRES may be used for the preparation of compositions for
XX the treatment of and/or prevention of a disease by gene therapy. Such
```

```
RESULT 4
AAA73710
ID AAA73710 standard; RNA; 222 BP.
XX
XX AAA73710;
XX
XX 07-DEC-2000 (first entry)
XX
XX RNA of human PITSLRE protein kinase gene internal ribosome entry site.
XX
XX IRES; internal ribosome entry site; PITSLRE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX misc_signal 1..222
XX /*tag= a
XX /label= IRES
XX /note= "Internal ribosome entry site"
XX
XX WO200044896-A1.
XX
XX 03-AUG-2000.
XX
XX 26-JAN-2000; 2000MO-EP00643.
XX
XX 26-JAN-1999; 99EP-0200216.
XX
XX (VIAA-) VIAAMS INTERUNIVERSITAIR INST BIOTECHNOC.
XX
XX Cornelis S, Beyaert R;
XX
XX WPI; 2000-499331/44.
XX
XX Nucleic acids encoding internal ribosome entry sequences useful for
XX directing protein expression in gene therapy procedures -
XX
XX Disclosure; Page 32; 57pp; English.
XX
XX The PITSLRE protein kinase gene can be translated to give two isoforms,
XX p110 and p58. Transcription of p58 mRNA was found to start at an
XX internal ribosome entry site (IRES). The IRES element was found to be
XX cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX cycle. The present sequence is the corresponding RNA of the PITSLRE
XX IRES described in AAA73710. The IRES sequence and fragments of it may be
XX used to induce cell cycle dependent initiation of translation in
XX eukaryotic cells. Vectors containing the IRES may be used for the
XX preparation of compositions for the treatment of and/or prevention of
XX a disease by gene therapy. Such diseases may be cancer and restenosis.
XX
XX Sequence 222 BP; 77 A; 36 C; 89 G; 20 U; 0 other:
XX
XX Query Match 100.0%; Score 30; DB 21; Length 222;
XX Best Local Similarity 96.7%; Pred. No. 0.0085;
XX Matches 29; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 gacatcagcagcagcgaggaagaccagc 30
XX ||||:|||||||||||||||||||||||
XX Db 1 gacaucagcagcagcgaggaagaccagc 30
XX
XX RESULT 5
XX AAA73713
XX ID AAA73713 standard; DNA; 468 BP.
XX
XX AAA73713;
XX
XX 07-DEC-2000 (first entry)
XX
XX 5' extension of PITSLRE protein kinase internal ribosome entry site.
XX
```

```
KW IRES; internal ribosome entry site; PITSLRE; human; gene therapy;
KW cancer; restenosis; p58; p110; protein kinase; ds.
XX
XX Homo sapiens.
XX
XX WO200044896-A1.
XX
XX 03-AUG-2000.
XX
XX 26-JAN-2000; 2000MO-EP00643.
XX
XX 26-JAN-1999; 99EP-0200216.
XX
XX (VIAA-) VIAAMS INTERUNIVERSITAIR INST BIOTECHNOC.
XX
XX Cornelis S, Beyaert R;
XX
XX WPI; 2000-499331/44.
XX
XX Nucleic acids encoding internal ribosome entry sequences useful for
XX directing protein expression in gene therapy procedures -
XX
XX Claim 9; Page 33; 57pp; English.
XX
XX The PITSLRE protein kinase gene can be translated to give two isoforms,
XX p110 and p58. Transcription of p58 mRNA was found to start at an
XX internal ribosome entry site (IRES). The IRES element was found to be
XX cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX cycle. The present sequence contains a 5' fragment of the PITSLRE
XX protein kinase IRES and nucleotides upstream of the IRES. The IRES
XX sequence and fragments of it may be used to induce cell cycle
XX dependent initiation of translation in eukaryotic cells. Vectors
XX containing the IRES may be used for the preparation of compositions for
XX the treatment of and/or prevention of a disease by gene therapy. Such
XX diseases may be cancer and restenosis.
XX
XX Sequence 468 BP; 144 A; 100 C; 182 G; 42 T; 0 other:
XX
XX Query Match 100.0%; Score 30; DB 21; Length 468;
XX Best Local Similarity 100.0%; Pred. No. 0.0091;
XX Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 gacatcagcagcagcgaggaagaccagc 30
XX |||||:|||||||||||||||||||||
XX Db 439 gacatcagcagcagcgaggaagaccagc 468
XX
XX RESULT 6
XX AAA73714
XX ID AAA73714 standard; DNA; 660 BP.
XX
XX AAA73714;
XX
XX 07-DEC-2000 (first entry)
XX
XX Human PITSLRE protein kinase gene IRES and upstream sequence.
XX
XX IRES; internal ribosome entry site; PITSLRE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX misc_signal 439..660
XX /*tag= a
XX /label= IRES
XX /note= "Internal ribosome entry site"
XX
XX WO200044896-A1.
XX
XX 03-AUG-2000.
XX
```

CC The PITSLRE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence is a 5' fragment of PITSLRE protein kinase
CC IRES. The IRES sequence and fragments of it may be used to induce cell
CC cycle dependent initiation of translation in eukaryotic cells. Vectors
CC containing the IRES may be used for the preparation of compositions for
CC the treatment of and/or prevention of a disease by gene therapy. Such
CC diseases may be cancer and restenosis.

XX
SQ Sequence 30 BP; 11 A; 8 C; 10 G; 1 T; 0 other;

Query Match 100.0%; Score 30; DB 21; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0069;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gacatcagcagcagcagcagcagcagcagc 30
|||||
DB 1 gacatcagcagcagcagcagcagcagcagc 30

RESULT 2

AA73728
ID AAA73728 standard; DNA; 40 BP.

XX
AC AAA73728;

XX
DT 07-DEC-2000 (first entry)

DE Sense primer di-6 used in creation of PITSLRE dicistronic vector.

XX IRES; internal ribosome entry site; PITSLRE; human; gene therapy;

KW cancer; restenosis; p58; p110; protein kinase; PCR primer;

KM ss.

OS Homo sapiens.

OS Synthetic.

PN WO200044896-A1.

XX
PD 03-AUG-2000.

XX
PF 26-JAN-2000; 2000WO-EP00643.

XX
PR 26-JAN-1999; 99EP-0200216.

XX
PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX
PI Cornelis S, Beyaert R;

XX
DR WPI: 2000-499331/44.

XX
PT Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -

XX
PS Examples: Page 25; 57pp; English.

XX
CC The PITSLRE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The IRES sequence and fragments of it may be used to induce cell
CC cycle dependent initiation of translation in eukaryotic cells. Vectors
CC containing the IRES may be used for the preparation of compositions for
CC the treatment of and/or prevention of a disease by gene therapy. Such
CC diseases may be cancer and restenosis. The present sequence is the
CC sense primer di-6 used in creation of a PITSLRE dicistronic vector.
CC This was used in the process of identifying the IRES.

XX
SQ Sequence 40 BP; 14 A; 10 C; 12 G; 4 T; 0 other;

Query Match 100.0%; Score 30; DB 21; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gacatcagcagcagcagcagcagcagcagc 30
|||||
DB 11 gacatcagcagcagcagcagcagcagcagc 40

RESULT 3

AA73709
ID AAA73709 standard; DNA; 222 BP.

XX
AC AAA73709;

XX
DT 07-DEC-2000 (first entry)

DE Human PITSLRE protein kinase gene internal ribosome entry site.

XX IRES; internal ribosome entry site; PITSLRE; human; gene therapy;

KW cancer; restenosis; p58; p110; protein kinase; ds.

XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers

FT misc_signal 1..222

FT /*tag= a

FT /label= IRES

FT /note= "internal ribosome entry site"

XX
PN WO200044896-A1.

XX
PD 03-AUG-2000.

XX
PF 26-JAN-2000; 2000WO-EP00643.

XX
PR 26-JAN-1999; 99EP-0200216.

XX
PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX
PI Cornelis S, Beyaert R;

XX
DR WPI: 2000-499331/44.

XX
PT Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -

XX
PS Claim 4; Page 32; 57pp; English.

XX
CC The PITSLRE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence is the PITSLRE protein kinase IRES. The
CC IRES sequence and fragments of it may be used to induce cell cycle
CC dependent initiation of translation in eukaryotic cells. Vectors
CC containing the IRES may be used for the preparation of compositions for
CC the treatment of and/or prevention of a disease by gene therapy. Such
CC diseases may be cancer and restenosis.

XX
SQ Sequence 222 BP; 77 A; 36 C; 89 G; 20 T; 0 other;

Query Match 100.0%; Score 30; DB 21; Length 222;
Best Local Similarity 100.0%; Pred. No. 0.0085;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 gacatcagcagcagcagcagcagcagcagc 30
|||||
DB 1 gacatcagcagcagcagcagcagcagcagc 30

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OK nucleic - nucleic search, using sw model

Run on: December 4, 2001, 19:03:35 : Search time 355.05 Seconds
(without alignments)
72.440 Million cell updates/sec

Title: US-09-915-060-4
Perfect score: 30
Sequence: 1 gacatcagcagacgagagagacacgac 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_1101: *
1: /SIDS2/gcgdata/geneseq/geneseq/NA1980.DAT: *
2: /SIDS2/gcgdata/geneseq/geneseq/NA1981.DAT: *
3: /SIDS2/gcgdata/geneseq/geneseq/NA1982.DAT: *
4: /SIDS2/gcgdata/geneseq/geneseq/NA1983.DAT: *
5: /SIDS2/gcgdata/geneseq/geneseq/NA1984.DAT: *
6: /SIDS2/gcgdata/geneseq/geneseq/NA1985.DAT: *
7: /SIDS2/gcgdata/geneseq/geneseq/NA1986.DAT: *
8: /SIDS2/gcgdata/geneseq/geneseq/NA1987.DAT: *
9: /SIDS2/gcgdata/geneseq/geneseq/NA1988.DAT: *
10: /SIDS2/gcgdata/geneseq/geneseq/NA1989.DAT: *
11: /SIDS2/gcgdata/geneseq/geneseq/NA1990.DAT: *
12: /SIDS2/gcgdata/geneseq/geneseq/NA1991.DAT: *
13: /SIDS2/gcgdata/geneseq/geneseq/NA1992.DAT: *
14: /SIDS2/gcgdata/geneseq/geneseq/NA1993.DAT: *
15: /SIDS2/gcgdata/geneseq/geneseq/NA1994.DAT: *
16: /SIDS2/gcgdata/geneseq/geneseq/NA1995.DAT: *
17: /SIDS2/gcgdata/geneseq/geneseq/NA1996.DAT: *
18: /SIDS2/gcgdata/geneseq/geneseq/NA1997.DAT: *
19: /SIDS2/gcgdata/geneseq/geneseq/NA1998.DAT: *
20: /SIDS2/gcgdata/geneseq/geneseq/NA1999.DAT: *
21: /SIDS2/gcgdata/geneseq/geneseq/NA2000.DAT: *
22: /SIDS2/gcgdata/geneseq/geneseq/NA2001.DAT: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	30	21	AAA73712
2	30	100.0	40	21	AAA73728
3	30	100.0	222	21	AAA73709
4	30	100.0	222	21	AAA73710
5	30	100.0	468	21	AAA73713
6	30	100.0	660	21	AAA73714
7	30	100.0	2471	21	AAA73711
8	28	93.3	35	21	AAA73726
9	20.6	68.7	522	22	AA118242
10	20.6	68.7	522	22	AA143284
11	20.6	68.7	1740	11	AA05747

C	12	20.6	68.7	1741	15	AA053994
C	13	20.6	68.7	1741	15	AA044854
C	14	20.6	68.7	1741	15	AA081891
C	15	20	66.7	831	21	AA049954
C	16	20	66.7	833	21	AA034562
C	17	20	66.7	1722	19	AAV20740
C	18	20	66.7	2443	19	AAV36468
C	19	20	66.7	2450	19	AAV20757
C	20	20	66.7	2918	21	AAZ43922
C	21	20	66.7	2919	21	AAZ24486
C	22	19.6	65.3	1261	21	AA067837
C	23	19.4	64.7	292	19	AAV62323
C	24	19.4	64.7	601	19	AAV40880
C	25	19.4	64.7	601	21	AA064148
C	26	19.4	64.7	601	22	AA098445
C	27	19.4	64.7	1128	21	AA030619
C	28	19.4	64.7	1128	21	AA030629
C	29	19.4	64.7	1128	21	AA030724
C	30	19.4	64.7	1128	21	AA030728
C	31	19.4	64.7	1648	21	AAZ11700
C	32	19.4	64.7	1841	19	AAV62310
C	33	19.4	64.7	2582	17	AA036034
C	34	19.4	64.7	2582	20	AA026245
C	35	19.4	64.7	2582	20	AAV81862
C	36	19.4	64.7	2582	21	AA088577
C	37	19.4	64.7	2582	21	AAZ89790
C	38	19.4	64.7	2582	21	AAZ88006
C	39	19.4	64.7	2776	18	AA073366
C	40	19.4	64.7	2894	19	AAV62334
C	41	19	63.3	182	14	AA042773
C	42	19	63.3	1077	21	AA021043
C	43	19	63.3	1077	21	AA034921
C	44	19	63.3	2296	19	AAV12457
C	45	19	63.3	2372	21	AA021041

ALIGNMENTS

RESULT	1
AA073712	
ID	AAA73712 standard; DNA; 30 BP.
XX	
AC	AAA73712;
XX	
DT	07-DEC-2000 (first entry)
XX	
DE	5' fragment of PITSLRE protein kinase internal ribosome entry site.
XX	
KW	IRRS; Internal ribosome entry site; PITSLRE; human; gene therapy;
KM	Cancer; restenosis; p58; p110; protein kinase; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200044896-A1.
XX	
PD	03-AUG-2000.
XX	
PF	26-JAN-2000; 2000WO-EP00643.
XX	
PR	26-JAN-1999; 99EP-0200216.
XX	
PA	(VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX	
PI	Cornellis S, Beyaert R;
XX	
DR	WPI; 2000-499331/44.
XX	
PT	Nucleic acids encoding internal ribosome entry sequences useful for
XX	directing protein expression in gene therapy procedures -
PS	Claim 8; Page 33; 57pp; English.
XX	

MEL EPO receptor c
Mouse soluble EPO
Mouse erythropoiet
Arabidopsis thalia
Arabidopsis thalia
Human squalene epo
Human squalene epo
Human squalene epo
S. tendae rtkkmyc
Streptomyces tend
Human betaf3 cDNA.
Human CS198 clone
5' portion of clon
Nucleotide sequenc
Human cDNA clone B
Human G protein-co
Human G protein-co
DNA encoding human
Human encoding human
Human CEPR (hCEPR)
Human CS198 genom
rchd523 gene diffe
cDNA sequence of r
Human rchd523 enco
Human rchd523 cDNA
Human rchd523 gene
rchd523 gene sequ
FGC-1 gene. Homo
Human CS198 consen
Ligand-induced gen
Human low adenosin
Human adenosine re
Human Hsp4 prostagl
Human low adenosin

FILE REFERENCE: 30062-20022.00 Polyketide Synthase Enz
CURRENT APPLICATION NUMBER: US/09/154,083
CURRENT FILING DATE: 1998-09-16
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 11
LENGTH: 750
TYPE: DNA
ORGANISM: Streptomyces hygroscopicus
US-09-154-083-11

Query Match 62.7%; Score 18.8; DB 1; Length 750;
Best Local Similarity 76.7%; Pred. No. 50;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 gacatcagcgacagcgagaggaagaccagc 30
|||||
DB 491 GACAGCAGCGACGTCGACGACGTCGACG 462

RESULT 15
US-08-222-289-1/c
Sequence 1, Application US/08222289
Patent No. 5559010
GENERAL INFORMATION:
APPLICANT: Kilonsky, Daniel J
APPLICANT: Destriele, Monika
TITLE OF INVENTION: NUTRIENT REGULATED GENE EXPRESSION
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,289
FILING DATE: 04-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59767/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2165 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 370..1434
US-08-222-289-1

Query Match 62.7%; Score 18.8; DB 1; Length 2165;
Best Local Similarity 76.7%; Pred. No. 55;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 gacatcagcgacagcgagaggaagaccagc 30

DB 1002 GTCATCAGCATAGCAACAGGCAACCCAGC 973
|||||

Search completed: December 4, 2001, 18:57:33
Job time: 9199 sec

CLONE: 10A8-T7
US-08-330-108-5

Query Match 63.3%; Score 19; DB 1; Length 182;
Best Local Similarity 81.5%; Pred. No. 37;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 catcagcgacgagcaggaagaccag 29
|| ||||| ||||| ||||| ||||| |||||
Db 32 CAGCAGCGCGCAGCAGCAGCAGCAGCAG 6

RESULT 12
PCT-US92-10087-5/C
; Sequence 5, Application PC/TUS9210087
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A.
; TITLE OF INVENTION: IL-2-Stimulated Gene
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: labive & Cockfield
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII TEXT
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/10087
; FILING DATE: 19921118
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/796,066
; FILING DATE: 20-NOVEMBER-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Decont, Giulio A. Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: DCI-028PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-227-7400
; TELEFAX: 617-227-5941
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 182 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single stranded
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: no
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: human
; CELL TYPE: T-cell blast
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: 10A8-T7
; PCT-US92-10087-5

Query Match 63.3%; Score 19; DB 5; Length 182;
Best Local Similarity 81.5%; Pred. No. 37;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 catcagcgacgagcaggaagaccag 29
|| ||||| ||||| ||||| ||||| |||||
Db 32 CAGCAGCGCGCAGCAGCAGCAGCAGCAG 6

RESULT 13
US-08-239-431A-3/C
; Sequence 3, Application US/08239431A
; Patent No. 5716835
; GENERAL INFORMATION:
; APPLICANT: Regan, John W.
; APPLICANT: Gil, Daniel W.
; APPLICANT: Woodward, David F.
; TITLE OF INVENTION: NOVEL HUMAN EP PROSTAGLANDIN RECEPTOR
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/239,431A
; FILING DATE: 05-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: ALRGN.053A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2296 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 157...1230
; OTHER INFORMATION:
; US-08-239-431A-3

Query Match 63.3%; Score 19; DB 1; Length 2296;
Best Local Similarity 81.5%; Pred. No. 46;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 catcagcgacgagcaggaagaccag 29
|| ||||| ||||| ||||| ||||| |||||
Db 675 CAGCAGCGCGCAGCAGCAGCAGCAGCAG 649

RESULT 14
US-09-154-083-11/C
; Sequence 11, Application US/09154083
; Patent No. 6150513
; GENERAL INFORMATION:
; APPLICANT: Wu, Kai
; TITLE OF INVENTION: Polypeptide Synthase Enzymes and Recombinant DNA
; CONSTRUCTS THEREFOR

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,496
FILING DATE: 06-OCT-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,654
FILING DATE: 09-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-104
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2582 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
US-08-944-496-6

Query Match 64.7%; Score 19.4; DB 3; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 2 acatcagcagcagcagcagcagcagc 30
Db 809 ACATCAGCTTCCGCGAGAGATGACCATC 837

RESULT 10
US-08-925-767-6
Sequence 6, Application US/08925767
Patent No. 6225084
GENERAL INFORMATION:
APPLICANT: FAIR, DEAN A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/925,767
FILING DATE: 09-SEPT-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844

FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-097
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2582 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
US-08-925-767-6

Query Match 64.7%; Score 19.4; DB 4; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 2 acatcagcagcagcagcagcagcagc 30
Db 809 ACATCAGCTTCCGCGAGAGATGACCATC 837

RESULT 11
US-08-330-108-5/C
Sequence 5, Application US/08330108
Patent No. 5795752
GENERAL INFORMATION:
APPLICANT: SMITH, Kendall A.
TITLE OF INVENTION: IL-2-Stimulated Gene
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII TEXT
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/330,108
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/104,736
FILING DATE:
APPLICATION NUMBER: US/07/796,066
FILING DATE:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 base pairs
TYPE: nucleic acid
STRANDEDNESS: single stranded
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: no
ANTI-SENSE: no
ORIGINAL SOURCE:
ORGANISM: human
CELL TYPE: T-cell blast
IMMEDIATE SOURCE:
LIBRARY:

Query Match 64.7%; Score 19.4; DB 3; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 2 acatcagcagcagcagcagcagcagc 30
||||| | | | | | | | | | | | | | | |
DB 809 ACATCAGCTCCGCGAGAGATGACCATC 837

RESULT 7
US-08-944-423A-6
; Sequence 6, Application US/08944423A
; Patent No. 6020463
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944.423A
; FILING DATE: 06-OCT-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,654
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: JUN-07-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-105
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
US-08-944-423A-6

Query Match 64.7%; Score 19.4; DB 3; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 2 acatcagcagcagcagcagcagcagc 30
||||| | | | | | | | | | | | | | | |
DB 809 ACATCAGCTCCGCGAGAGATGACCATC 837

RESULT 8
US-08-925-743-6

; Sequence 6, Application US/08925743
; Patent No. 6054558
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/925,743
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/485,573
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
US-08-925-743-6

Query Match 64.7%; Score 19.4; DB 3; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 2 acatcagcagcagcagcagcagcagc 30
||||| | | | | | | | | | | | | | | |
DB 809 ACATCAGCTCCGCGAGAGATGACCATC 837

RESULT 9
US-08-944-496-6
; Sequence 6, Application US/08944496
; Patent No. 6124433
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; SOFTWARE: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/599,654
FILING DATE: 09-FEB-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-041
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2582 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
US-08-599-654-6

Query Match 64.7%; Score 19.4; DB 2; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 acatcagcgacgcgaggaagaccagc 30
DB 809 ACATCAGCTTCGCGAGAGATGACCATC 837

RESULT 5
US-08-485-573-6
Sequence 6, Application US/08485573
Patent No. 5968770
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,573
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-032
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2582 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
US-08-485-573-6

Query Match 64.7%; Score 19.4; DB 2; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 acatcagcgacgcgaggaagaccagc 30
DB 809 ACATCAGCTTCGCGAGAGATGACCATC 837

RESULT 6
US-08-944-868A-6
Sequence 6, Application US/08944868A
Patent No. 6018025
GENERAL INFORMATION:
APPLICANT: FALB, DEAN A
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,868A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/599,654
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-041
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2582 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
US-08-944-868A-6

QY 3 catcagcagcagcagagagaccagc 30
||| ||||| ||||| |||||
DB 799 CACCGAGCCGACGAGAGACACGAGC 772

RESULT 2
US-08-480-994-6
; Sequence 6, Application US/08480994
; Patent No. 5834248
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,994
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; US-08-480-994-6

Query Match 64.7%; Score 19.4; DB 2; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 acatcagcagcagcagagagaccagc 30
||||| ||||| ||||| |||||
DB 809 ACATCAGCTTCGCGAGAGATGACCATC 837

RESULT 3
US-08-616-844-6
; Sequence 6, Application US/08616844
; Patent No. 5849578
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A.
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE

NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/616,844
FILING DATE: 15-MAR-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/599,654
FILING DATE: 09-FEB-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,573
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2582 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
US-08-616-844-6

Query Match 64.7%; Score 19.4; DB 2; Length 2582;
Best Local Similarity 79.3%; Pred. No. 33;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 acatcagcagcagcagagagaccagc 30
||||| ||||| ||||| |||||
DB 809 ACATCAGCTTCGCGAGAGATGACCATC 837

RESULT 4
US-08-599-654-6
; Sequence 6, Application US/08599654
; Patent No. 5882925
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:57:28 ; Search time 152.61 seconds
(without alignments)
44.521 Million cell updates/sec

Title: US-09-915-060-4

Perfect score: 30

Sequence: 1 gacatcagcagcagcagcagcagcagcagc 30

Scoring table: IDENTITY_NUC

Searched: Gap0 10.0 , Gapext 1.0

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	66.7	2443	2	US-08-745-934-2
2	19.4	64.7	2582	2	US-08-480-994-6
3	19.4	64.7	2582	2	US-08-616-844-6
4	19.4	64.7	2582	2	US-08-599-654-6
5	19.4	64.7	2582	2	US-08-485-573-6
6	19.4	64.7	2582	3	US-08-944-868A-6
7	19.4	64.7	2582	3	US-08-944-423A-6
8	19.4	64.7	2582	3	US-08-925-743-6
9	19.4	64.7	2582	3	US-08-944-496-6
10	19.4	64.7	2582	4	US-08-925-767-6
11	19	63.3	182	5	US-08-330-108-5
12	19	63.3	182	5	US-08-330-108-5
13	19	63.3	182	5	US-08-330-108-5
14	18.8	62.7	750	1	US-09-154-083-11
15	18.8	62.7	750	1	US-08-222-289-1
16	18.4	61.3	719	3	US-08-983-409-6
17	18.4	61.3	1696	1	US-07-887-072B-3
18	18.4	61.3	1696	1	US-08-466-444-3
19	18.4	61.3	1884	3	US-08-753-007A-5
20	18.4	61.3	1884	3	US-09-398-496-5
21	18.4	61.3	3224	3	US-09-079-415-3
22	17.8	59.3	70	4	US-09-210-016-2
23	17.8	59.3	111	1	US-08-454-720A-42
24	17.8	59.3	992	1	US-08-358-782D-13
25	17.8	59.3	992	2	US-08-764-527A-13
26	17.8	59.3	1462	1	US-08-358-782D-14
27	17.8	59.3	1462	2	US-08-764-527A-14

28	17.8	59.3	1549	2	US-08-865-597A-1	Sequence 1, Appli
29	17.8	59.3	2122	4	US-09-029-603-1	Sequence 1, Appli
30	17.8	59.3	7130	4	US-09-056-105-31	Sequence 31, Appli
31	17.8	59.3	9551	1	US-08-056-200-93	Sequence 93, Appli
32	17.8	59.3	9551	2	US-08-800-644-93	Sequence 93, Appli
33	17.8	59.3	4403765	4	US-09-103-840A-2	Sequence 2, Appli
34	17.8	59.3	4411529	4	US-09-103-840A-1	Sequence 1, Appli
35	17.4	58.0	333	2	US-08-634-797-21	Sequence 21, Appli
36	17.4	58.0	1074	2	US-08-463-081B-29	Sequence 29, Appli
37	17.4	58.0	1074	2	US-08-461-379A-29	Sequence 29, Appli
38	17.4	58.0	1074	3	US-08-462-390B-29	Sequence 29, Appli
39	17.4	58.0	1074	3	US-08-463-074B-29	Sequence 29, Appli
40	17.4	58.0	1074	3	US-08-465-585C-29	Sequence 29, Appli
41	17.4	58.0	1074	3	US-08-652-446-29	Sequence 29, Appli
42	17.4	58.0	2450	2	US-08-463-081B-5	Sequence 29, Appli
43	17.4	58.0	2450	2	US-08-461-379A-5	Sequence 5, Appli
44	17.4	58.0	2450	2	US-08-462-390B-5	Sequence 5, Appli
45	17.4	58.0	2450	3	US-08-463-074B-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-745-934-2/c
; Sequence 2, Application US/08745934
; Patent No. 5861496
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: HUMAN SQUALENE EPOXIDASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/745,934
; FILING DATE: Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0151 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 855-0555
; TELEFAX: (415) 845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2443 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE: Consensus
; US-08-745-934-2

Query Match 66.7%; Score 20; DB 2; Length 2443;
Best Local Similarity 82.1%; Pred. No. 19;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Query Match 86.0%; Score 25.8; DB 10; Length 216;
Best Local Similarity 93.1%; Pred. No. 14;
Matches 27; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 gacatcagcgacagcgaggaagaccag 29
||||||| |||||||
Db 86 GACATCAGTCAGCAGCGAGGAAAAACCAG 114

Search completed: December 4, 2001, 18:14:19
Job time: 6760 sec

ORIGIN

Query Match 89.3%; Score 26.8; DB 10; Length 649;
Best Local Similarity 93.3%; Pred. No. 7.3;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gacatcagcgacagcgaggaagaccagc 30
||||||| ||||||| ||||||| |||||||
Db 44 GACATCAGTGCAGCGAGGAGAAACCGC 73

RESULT 12

LOCUS BG298479 900 bp mRNA EST 21-FEB-2001
DEFINITION 602396747F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4511434 5',
mRNA sequence.
ACCESSION BG298479.1 GI:13063173
VERSION BG298479
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 900)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rt@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM10394 row: n column: 11
High quality sequence stop: 676.
Location/Qualifiers
1..900
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_image="4511434"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."

BASE COUNT 302 a 175 c 300 g 123 t

ORIGIN

Query Match 89.3%; Score 26.8; DB 11; Length 900;
Best Local Similarity 93.3%; Pred. No. 7.7;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gacatcagcgacagcgaggaagaccagc 30
||||||| ||||||| ||||||| |||||||
Db 47 GACATCAGTGCAGCGAGGAGAAACCGC 76

RESULT 13

LOCUS CNS0431P/C 972 bp DNA GSS 18-MAY-2000
DEFINITION Tetradon nigroviridis genome survey sequence PUC-ori end of clone
079D08 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL272842
VERSION AL272842.1 GI:7995084

KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigroviridis.
ORGANISM Tetradon nigroviridis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetradon.
REFERENCE 1 (bases 1 to 972)
AUTHORS Roest-Collins/H., Jallion/O., Dasilva/C., Fitzames/C., Fisher/C.,
Bonneau/L., Billault/A., Quetier/F., Saurin/W., Bernot/A. and
Weissenbach/J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 972)
AUTHORS Roest-Collins/H., Jallion/O., Dasilva/C., Bonneau/L., Fisher/C.,
Bernot/A., Fitzames/C., Wincker/P., Brothier/P., Quetier/F.,
Saurin/W. and Weissenbach/J.
TITLE Human gene number estimate provided by genome wide analysis using
Tetradon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 972)
AUTHORS Genoscope.
TITLE Direct Submission
COMMENT Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-and-sequencing project of the Tetradon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetradon>.
Location/Qualifiers
1..972
/organism="Tetradon nigroviridis"
/db_xref="taxon:99883"
/clone_image="079D08"
/clone_lib="G"
/note="Genoscope sequence ID : C0BG079DB04SP1-end :
PUC-ori"

BASE COUNT 224 a 249 c 220 g 272 t 7 others

ORIGIN

Query Match 89.3%; Score 26.8; DB 13; Length 972;
Best Local Similarity 93.3%; Pred. No. 7.8;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 gacatcagcgacagcgaggaagaccagc 30
||||||| ||||||| ||||||| |||||||
Db 696 GACATCAGTGCAGCGAGGAGAAACCGC 667

RESULT 14

LOCUS AK008283 1000 bp mRNA HTC 05-JUL-2001
DEFINITION Mus musculus adult male small intestine cDNA, RIKEN full-length
enriched library, clone:2010016J05, full insert sequence.
ACCESSION AK008283
VERSION AK008283.1 GI:12842370
KEYWORDS CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) adult male small intestine cDNA to
mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
clone:2010016J05.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1000)
AUTHORS Carninci/P. and Hayashizaki/Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Methods in enzymology. 303, 19-44 (1999)
MEDLINE 99279253
PubMed 10349636
REFERENCE 2 (bases 1 to 1000)
AUTHORS Carninci/P., Shibata/Y., Hayatsu/N., Sugahara/Y., Shibata/K.,
Itoh/M., Konno/H., Okazaki/Y., Muramatsu/M. and Hayashizaki/Y.

Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:490063

Putative full length read
vector to vector length is 183
Seq primer: -28ml3 rev1 ET from Amersham.

FEATURES

SOURCE

Location/Qualifiers

1. 182

/organism="Mus musculus"

/strain="C57BL/6 x DBA"

/db_xref="taxon:10090"

/clone="IMAGE:821783"

/clone_1lb="Beddington mouse embryonic region"

/sex="pooled"

/tissue_type="embryo"

/dev_stage="7.5dpc"

/lab_host="DH12S"

/note="Organ: whole embryo; Vector: PCMV-SPORT; Site:1;

SalI; Site:2: NotI; Cloned unidirectionally. Primer:

Oligo dT. Gastrulating embryos were collected at 7.5dpc

from C57BL/6 x DBA matings, excluding embryos that had

developed head folds and all extraembryonic tissues.

Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).

Referenced in Development 121, 2479-2489 (1995)"

BASE COUNT

66 a 32 c 59 g 25 t

ORIGIN

Query Match

Best Local Similarity 93.3%; Score 26.8; DB 10; Length 182;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagcagc 30
Db 88 GACATCAGTCAGCAGCAGCAGCAGCAGCAGC 117

RESULT 10

AL034747

LOCUS AL034747 372 bp mRNA EST 29-DEC-1999

DEFINITION V6608b53 Beddington mouse dissected endoderm Mus musculus CDNA

clone 528_8K10 5', mRNA sequence.

AL034747

AL034747.1 GI:6646373

EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 372)

Harrison,S.M., Dunwoodie,S.L., Arkell,R.M., Lehrach,H. and

Beddington,R.S.

Isolation of novel tissue-specific genes from cDNA libraries

representing the individual tissue constituents of the gastrulating

mouse embryo

Development 121 (8), 2479-2489 (1995)

95401865

CONTACT: Wiles,M., Lehrach,H. and Avner,P.

ERC Mouse Transcript Mapping Consortium

Genoscope - CNS

2, rue Gaston Cremieux, 91000 Evry, France

Email: payner@pasteur.fr

clone available from Ressourcenzentrum, Heubnerweg 6, D-14059

Berlin, Germany. Web site http://www.rzpd.de

Seq primer: CCGGTCGGGATTCGCGG

High quality sequence only submitted.

Vector: pSPORT; site:1: NotI; site:2: SalI;

Cloned unidirectionally.

Dissected endoderm 7.5 days.

Average insert size: 1.2 kb (range: 0.2 - 2.kb).

Location/Qualifiers

1. .372

FEATURES

source

BASE COUNT

126 a 74 c 130 g 42 t

ORIGIN

Query Match

Best Local Similarity 93.3%; Score 26.8; DB 10; Length 372;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagcagc 30
Db 266 GACATCAGTCAGCAGCAGCAGCAGCAGCAGC 295

RESULT 11

AW210026

LOCUS AW210026 649 bp mRNA EST 03-DEC-1999

DEFINITION u150a11.y1 Rasbass mouse MOV 9 5 optic vesicle Mus musculus CDNA

clone IMAGE:2101724 5', similar to gb:U37092 Mus musculus

cyclin-dependent kinase homologue (MOUSE);, mRNA sequence.

AW210026

AW210026.1 GI:6515966

EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 649)

Marrin,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,

Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person

,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Rittler

,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,

Waterston,R. and Wilson,R.

The WashU-NCI Mouse EST Project 1999

Unpublished (1999)

CONTACT: Marrin M/WashU-NCI Mouse EST Project 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.lnl.gov) for further information.

Seq primer: 998656

MGI:998656

High quality sequence stop: 426.

Location/Qualifiers

1. .649

/organism="Mus musculus"

/strain="CD1"

/db_xref="taxon:10090"

/clone="IMAGE:2101724"

/clone_1lb="Rasbass mouse MOV 9 5 optic vesicle"

/tissue_type="Optic vesicle and lens placode"

/dev_stage="embryo - 9.5 dpc"

/lab_host="DH10B"

/note="Organ: eye; Vector: pSPORT1; mRNA made from

developing eye tissue. cDNA made by oligo-dT priming with

NotI oligo. SalI adaptor (5'-TCGACCCAGCGTCG-3')

ligated to 5' ends. Size-selected with cDNA size

fractionation resin, average insert size 1.3 kb. Primary

library, non-amplified. Library constructed by Dr. Pen

Rasbass (pen@hu.mrc.ac.uk).

BASE COUNT

229 a 111 c 207 g 102 t

Query Match 100.0%; Score 30; DB 10; Length 1042;
 Best Local Similarity 100.0%; Pred. No. 0.6;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagacgagagaagaccagc 30
 ||||||||||||||||||||||||||||
 Db 413 GACATCAGCAGCAGCAGAGAGACCAGC 442

RESULT 7
 BE740559 1063 bp mRNA EST 15-SEP-2000
 LOCUS 601595656p1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3949693 5',
 DEFINITION mRNA sequence.
 BE740559
 VERSION BE740559.1 GI:10154551
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1063)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: L1CM814 row: P column: 14
 High quality sequence stop: 741.
 Location/Qualifiers
 1. 1063
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3949693"
 /clone_1lb="NIH_MGC_9"
 /tissue_type="adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: ovary; Vector: pORF7; Site: 1; XhoI; Site_2:
 EcoRI; CDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 351 a 203 c 402 g 107 t

ORIGIN

Query Match 100.0%; Score 30; DB 10; Length 1063;
 Best Local Similarity 100.0%; Pred. No. 0.6;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagacgagagaagaccagc 30
 ||||||||||||||||||||||||||||
 Db 408 GACATCAGCAGCAGCAGAGACCAGC 437

RESULT 8
 BG107702 1340 bp mRNA EST 30-JAN-2001
 LOCUS 602277878f1 NIH_MGC_86 Homo sapiens CDNA clone IMAGE:4365474 5',
 DEFINITION mRNA sequence.
 BG107702
 ACCESSION BG107702.1 GI:12601548
 VERSION

KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1340)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM10014 row: 1 column: 19
 High quality sequence stop: 667.
 Location/Qualifiers
 1. 1340
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4365474"
 /clone_1lb="NIH_MGC_86"
 /tissue_type="osteosarcoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: bone; Vector: pCMV-SPORT6; Site: 1; NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.533 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC library."

BASE COUNT 467 a 272 c 450 g 151 t

ORIGIN

Query Match 100.0%; Score 30; DB 11; Length 1340;
 Best Local Similarity 100.0%; Pred. No. 0.63;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagacgagagaagaccagc 30
 ||||||||||||||||||||||||||||
 Db 276 GACATCAGCAGCAGCAGAGACCAGC 305

RESULT 9
 AA413247 182 bp mRNA EST 07-MAY-1997
 LOCUS AA413247
 DEFINITION ve52d12.f1 Beddington mouse embryonic region Mus musculus CDNA
 clone IMAGE:821783 5' similar to gb:M58633 Mouse p58/GTA protein
 kinase mRNA, complete cds (MOUSE);, mRNA sequence.

ACCESSION AA413247
 VERSION AA413247.1 GI:2071810
 EST.

KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Scurionathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 182)
 AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
 Geisler, S., Kucada, T., Lacy, M., Le, M., Martin, J., Morris, B.,
 Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
 Thaising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
 Waterston, R.
 TITLE The WashU-HMI Mouse EST Project
 JOURNAL Unpublished (1996)
 COMMENT Contact: Marra M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

```

ACCESSION      BE742222
VERSION        BE742222.1  GI:10156214
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE         NIH-MGC http://mgc.ncl.nih.gov/.
JOURNAL       National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT       Unpublished (1999)
               Contact: Robert Strausberg, Ph.D.
               Email: cgabbs-remail.nih.gov
               Tissue Procurement: DCTD/DTP
               CDNA Library Preparation: Ling Hong/Rubin Laboratory
               CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
               Plate: L1CM520 row: j column: 08
               High quality sequence stop: 702.
               Location/Qualifiers
                 1..723
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone_image="3836647"
                   /clone_lib="NIH-MGC_9"
                   /tissue_type="adenocarcinoma cell line"
                   /lab_host="DH10B (phage-resistant)"
                   /note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
                   EcoRI; CDNA made by oligo-dT priming. Directionally
                   cloned into EcoRI/XhoI sites using the following 5'
                   adaptor: GGCAAGAG(G). Size-selected >500bp for average
                   insert size 1.8kb. Library constructed by Ling Hong in
                   the laboratory of Gerald M. Rubin (University of
                   California, Berkeley) using ZAP-cDNA synthesis kit
                   (Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT      230 a 140 c 284 g 69 t
ORIGIN
```

```

Query Match      100.0%; Score 30; DB 10; Length 723;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gacatcagcagcagcagcagcagcagcagc 30
Db 407 GACATCAGCAGCAGCAGCAGCAGCAGCAGC 436
```

```

RESULT 5
LOCUS      AL043340      740 bp      mRNA      EST      29-FEB-2000
DEFINITION DKEP43400923_r1.434 (synonym: htes3) Homo sapiens CDNA clone
ACCESSION  AL043340
VERSION     AL043340.1  GI:5422730
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE       1 (bases 1 to 740)
JOURNAL     Blum, H., Bauersachs, S., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
COMMENT     EST (Blum, et al.)
             Unpublished (1999)
             Contact: Blum H
             MIPS
             Am Klopferspitz 18a D-82152 Martinsried, Germany
             This is the 5' sequence of the clone insert
             Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
             Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
             sequenced by LMU (Ludwig Maximilians University,
```

Munich/Germany) within the CDNA sequencing consortium of the German Genome Project. No st sequence available. This clone (DKEP43400923) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

```

FEATURES
source
1..740
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /clone_image="DKEP43400923"
  /clone_lib="434 (synonym: htes3)"
  /tissue_type="testis"
  /dev_stage="adult"
  /lab_host="DH10B"
  /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
```

```

BASE COUNT      230 a 155 c 254 g 100 t 1 others
ORIGIN
```

```

Query Match      100.0%; Score 30; DB 10; Length 740;
Best Local Similarity 100.0%; Pred. No. 0.57;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gacatcagcagcagcagcagcagcagcagc 30
Db 44 GACATCAGCAGCAGCAGCAGCAGCAGCAGC 73
```

```

RESULT 6
LOCUS      BE742943      1042 bp      mRNA      EST      15-SEP-2000
DEFINITION 601574925P1 NIH-MGC_9 Homo sapiens CDNA clone IMAGE:3835686 5',
MRNA sequence.
ACCESSION    BE742943
VERSION      BE742943.1  GI:10157033
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE        1 (bases 1 to 1042)
JOURNAL      NIH-MGC http://mgc.ncl.nih.gov/.
COMMENT      National Institutes of Health, Mammalian Gene Collection (MGC)
               Unpublished (1999)
               Contact: Robert Strausberg, Ph.D.
               Email: cgabbs-remail.nih.gov
               Tissue Procurement: DCTD/DTP
               CDNA Library Preparation: Ling Hong/Rubin Laboratory
               CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
               Plate: L1CM518 row: b column: 07
               High quality sequence stop: 667.
               Location/Qualifiers
                 1..1042
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone_image="3835686"
                   /clone_lib="NIH-MGC_9"
                   /tissue_type="adenocarcinoma cell line"
                   /lab_host="DH10B (phage-resistant)"
                   /note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:
                   EcoRI; CDNA made by oligo-dT priming. Directionally
                   cloned into EcoRI/XhoI sites using the following 5'
                   adaptor: GGCAAGAG(G). Size-selected >500bp for average
                   insert size 1.8kb. Library constructed by Ling Hong in
                   the laboratory of Gerald M. Rubin (University of
                   California, Berkeley) using ZAP-cDNA synthesis kit
                   (Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT      371 a 159 c 390 g 122 t
ORIGIN
```

/note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

BASE COUNT 118 a 57 c 131 g 39 t

ORIGIN

Query Match 100.0%; Score 30; DB 11; Length 345;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagc 30
|||||

DB 13 GACATCAGCGACGAGAGAGACCAGC 42

RESULT 2
BF156008 516 bp mRNA EST 30-OCT-2000
LOCUS BF156008
DEFINITION RC0-HT0955-270900-033-d08 HT0955 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF156008
VERSION BF156008.1 GI:11051191
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 516)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R., Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=4t2-RC0-HT0955-270900-033-d08&ts=2000-09-27&td=1)
Seq primer: puc 18 forward
High quality sequence start: 56
High quality sequence stop: 516.
Location/Qualifiers
1. 516
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0955"
/dev_stage="Adult"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 155 a 109 c 175 g 77 t

ORIGIN

Query Match 100.0%; Score 30; DB 11; Length 516;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagc 30
|||||

DB 245 GACATCAGCGACGAGAGAGAGACCAGC 274

RESULT 3
BG822824 662 bp mRNA EST 22-MAY-2001
LOCUS BG822824
DEFINITION 602727816r1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4867443 5', mRNA sequence.
ACCESSION BG822824
VERSION BG822824.1 GI:14170411
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 662)
NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: NIH Intramural Sequencing Center
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L10M1735 row: p column: 04
High quality sequence stop: 601.
Location/Qualifiers
1. 662
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4867443"
/clone_lib="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies) "

FEATURES
source

BASE COUNT 202 a 139 c 255 g 66 t

ORIGIN

Query Match 100.0%; Score 30; DB 11; Length 662;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagc 30
|||||

DB 253 GACATCAGCGACGAGAGAGACCAGC 282

RESULT 4
BE742222 723 bp mRNA EST 15-SEP-2000
LOCUS BE742222
DEFINITION 601575629F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3836647 5', mRNA sequence.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:14:18 ; Search time 3881.49 Seconds
(without alignments)
83.054 Million cell updates/sec

Title: US-09-915-060-4

Sequence: 1 gacatcagcagcagcagcagcagcagcagc 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST: *
1: em_estfun: *
2: em_esthum: *
3: em_estin: *
4: em_estom: *
5: em_estpl: *
6: em_estba: *
7: em_estro: *
8: em_estov: *
9: em_hic: *
10: qb_estl: *
11: qb_est2: *
12: qb_hic: *
13: qb_gss: *
14: em_gss_fun: *
15: em_gss_hum: *
16: em_gss_inv: *
17: em_gss_pln: *
18: em_gss_pro: *
19: em_gss_rtd: *
20: em_gss_vrt: *
21: em_gss_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	345	11	BI261348 602953350
2	30	100.0	516	11	BF156008 RC0-HT095
3	30	100.0	662	11	BG822824 602727816
4	30	100.0	723	10	BE742222 601575629
5	30	100.0	740	10	AL043340 DKEP24340
6	30	100.0	1042	10	BE742943 601574925
7	30	100.0	1063	10	BE740559 601595656
8	30	100.0	1340	11	BG107702 602277878
9	26.8	89.3	182	10	AA413247 v652612.1
10	26.8	89.3	372	10	AL034747 v8608b53
11	26.8	89.3	649	10	AW210026 u150a11.y
12	26.8	89.3	900	11	BG298479 602396747

c	13	26.8	89.3	972	13	CNS0431P	AL272842 Tetradon
	14	26.8	89.3	1000	12	AK008283	AK008283 Mus muscu
	15	25.8	86.0	216	10	AV105652	AV105652 AV105652
	16	25.8	86.0	251	10	AV270235	AV270235 AV270235
	17	25.8	86.0	253	10	AV327502	AV327502 AV327502
	18	25.2	84.0	233	10	BB290670	BB290670 BB290670
	19	25.2	84.0	367	10	AA681955	AA681955 vt44h10.s
	20	25.2	84.0	493	11	BI018825	BI018825 IL3-MT026
	21	25.2	84.0	758	11	BG669359	BG669359 602836841
	22	25.2	84.0	812	11	BI149024	BI149024 602911290
	23	25.2	84.0	830	11	BF607177	BF607177 M1_00011
	24	24.2	80.7	147	10	AV065528	AV065528 AV065528
	25	24.2	80.7	192	10	AV368285	AV368285 AV368285
	26	22.6	75.3	655	13	CNS01FLJ	AL141992 Anopheles
	27	22.6	75.3	651	13	CNS01K2L	AL147806 Anopheles
	28	22.6	75.3	681	13	CNS01GSW	AL143553 Anopheles
	29	22.6	75.3	685	13	CNS01M6Q	AL150547 Anopheles
	30	22.6	75.3	721	13	CNS01F9Q	AL141567 Anopheles
	31	22.6	75.3	750	13	CNS01FRW	AL142221 Anopheles
	32	22.6	75.3	762	10	AJ281330	AJ281330 4A3A-P2A8
	33	22.6	75.3	788	13	CNS01EBU	AL140239 Anopheles
	34	22.6	75.3	1011	13	CNS01FEW	AL143581 Anopheles
	35	22.6	75.3	1024	13	CNS01F6S	AL141461 Anopheles
	36	22.6	75.3	1151	11	BF140125	BF140125 601786074
	37	22	73.3	566	13	AQ400706	AQ400706 HS_5069_A
	38	21.2	70.7	985	13	CNS05T01	AL353522 Tetradon
	39	21.2	70.7	1050	13	CNS04UP5	AL308066 Tetradon
	40	21	70.0	415	10	AI515282	AI515282 LD47038.5
	41	21	70.0	513	10	AA390276	AA390276 LD08974.5
	42	21	70.0	547	13	AO012093	AO012093 575P1H107
	43	21	70.0	570	10	AA951266	AA951266 LD31649.5
	44	21	70.0	571	10	AA979394	AA979394 LD33942.5
	45	21	70.0	586	11	BG985824	BG985824 515513 NI

ALIGNMENTS

RESULT 1
BI261348 LOCUS 345 bp mRNA EST 17-JUL-2001
602953350F1 NIH_MGC_99 Homo sapiens CDNA IMAGE:5087556 5',
DEFINITION mRNA sequence.
ACCESSION BI261348
VERSION BI261348.1 GI:14820527
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
1 (bases 1 to 345)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-r@mail.nih.gov
Tissue Procurement: Dr. Louis Staudt, M.D., Ph.D. CDNA library
Preparation: Ling Hong/Rubin laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM1845 row: c column: 13
High quality sequence stop: 345.
Location/Qualifiers

FEATURES
source
1..345
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5087556"
/clone_11b="NIH_MGC_99"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"

	gene	/db_xref="taxon:9606" /chromosome="11" /map="lp36.3" 1..2323 /gene="CDC2L2" 1104..2297 /gene="CDC2L2"
CDS		/note="corresponding genomic sequence deposited as GenBank Accession Numbers AF080689-AF080697 and AF092426-AF092428"
BASE COUNT	673 a 550 c 754 g 346 t	
ORIGIN		
Query Match	100.0%; Score 30; DB 9; Length 2323;	
Best Local Similarity	100.0%; Pred. No. 0.28;	
Matches	30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 gacatcagcgacagcgaagaaccacgc 30 Db 875 GACATCAGCGACGACGAGCAAGACCACGC 904	
RESULT 12		
AF067525	2329 bp mRNA PRI 07-NOV-1998	
LOCUS	AF067525 Homo sapiens PITSLR protein kinase beta SV13 isoform (CDC2L2)	
DEFINITION	mRNA, complete cds.	
VERSION	AF067525	
KEYWORDS	AF067525.1 GI:3850329	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2329) Gururajan,R., Lahit,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P.F. and Kidd,V.J. Duplication of a genomic region containing the Cdc2L1-2 and MP21-22 genes on human chromosome Ip36.3 and their linkage to D1Z2 Genome Res. 8 (9), 929-939 (1998)	
TITLE	2 (bases 1 to 2329) Gururajan,R., Lahit,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P. and Kidd,V.J. Direct Submission Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA	
JOURNAL MEDLINE	Location/Qualifiers	
AUTHORS	1..2329	
FEATURES	/organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="11" /map="lp36.3" 1..2329 /gene="CDC2L2" 68..1654 /note="Corresponding genomic sequence deposited as Genbank Accession Numbers AF080689-AF080697 and AF092426-AF092428" /codon_start=1 /product="PITSLR protein kinase beta SV13 isoform"	
source		

Query Match	Best Local Similarity	100.0%	Score 30;	DB 9;	Length 2329;
Matches 30;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY 1	gacatcagcgacagcgagaagaccagc 30				
Db 689	GACATCAGCGACAGCGAGGAAGCACGAC 718				
RESULT 13	HSU04818	2340 bp	mRNA	PII	08-JUL-1994
LOCUS	Human protein kinase PITSURE alpha 2-4	mRNA, complete cds.			
DEFINITION	004818				
ACCESSION	004818				
VERSION	004818.1	GI:507163			
KEYWORDS	human.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota: Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia: Eutheria: Primates; Catarrhini; Hominiidae: Homo.				
TITLE	1 (bases 1 to 2340)				
JOURNAL	Xiang, J., Lahli, J.M., Grenet, J.A., Easton, J.B. and Kidd, V.J.				
MEDLINE	Molecular cloning and expression of alternatively spliced PITSURE				
REFERENCE	protein kinase isoforms				
AUTHORS	J. Biol. Chem. 269, 15786-15794 (1994)				
TITLE	94253170				
JOURNAL	2 (bases 1 to 2340)				
REFERENCE	Kidd, V.J.				
AUTHORS	Direct Submission				
TITLE	Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's				
JOURNAL	Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St.,				
REFERENCE	Memphis, TN 38101, USA				
FEATURES	Location/Qualifiers				
SOURCE	1..2340				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/map="1p36-2"				
	/cell_line="Hela S3"				
	/cell_type="epitheloid"				
	/tissue_type="cervix"				
	/clone_lib="Hela CDNA library"				
	/dev_stage="adult"				
	1..2340				
	626..2314				
	/codon_start=1				
	/product="PITSURE alpha 2-4"				
	/protein_id="AA19584.1"				
	/db_xref="GI:507164"				
	/translation="MREDYADKYKASMSRSPRPREFELGDGKRPVKEKKEED				
	LLSLDQIDSDERKTSASSSASGSGSEEEEEEDEEGSTSESEEEEEEE				
	ETGSSSEASQSAEEDSVESSEEDERENMLLVPSRFDRDGSSEAEVEGEE				
	GPSPASLTEDGVYDSDPALSPTELKQELPKYLPALOGCRVSEFQCLNREGTG				
	VYRAADKDEIVALKRLKMEKEGEPITSLSAEINTLKAOHPNIVTRETIVG				
	SNMDDKIVYVNHDLKSLMETMKOPLPGEVKTLMQLRGVYKHLDMNILLRDLKTSN				
	LLISAGLIKRGDGLRGAIRGKSPKAYTPVYVYTVYRAPELLGAKKEYSTAVDMWSVG				
	CITGELTLQRLPFGKSEIDQIKVFRKRDITPSEKLTWDSGSLPVAWKSRHPYNN				
	LRKRGALLSDGQFLDNKFLTYFPERRISAEQGLKHEYFRFTPLRIDPSMFPMPAPK				

LOCUS AF174497 1715 bp mRNA PRI 29-FEB-2000
DEFINITION Homo sapiens GR AF-1 coactivator 3 mRNA, partial cds.
ACCESSION AF174497
VERSION AF174497.1 GI:7108916
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Govindan,M.V. and Warrier,N.
TITLE Human GR AF-1 Coactivator 3
JOURNAL Unpublished
AUTHORS Govindan,M.V.
TITLE 2 (bases 1 to 1715)
REFERENCE Direct Submission
AUTHORS Submitted (29-Jul-1999) Cancer Research, CRHD U. Laval, 9 Rue
JOURNAL McMahon, quebec, PQ G1R2U6, Canada
FEATURES
source Location/Qualifiers
1..1715
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="placenta"
/note="Isolated by interaction screening with GST-GR AF-1"
<..1699
/note="GR specific protein kinase"
/codon_start=2
/product="GR AF-1 coactivator 3"
/protein_id="AAF36538.1"
/db_xref="GI:7108917"
/translation="NSGMRKEDYSDKVKASHSRSPRPREFELGGRKPKKEKME
ENDLISDQIDISERKTSASSESSAEGSGSESEEESEEGSGTTESEEEEEE
EEEGTSNSEESESEAEVESESESESESESESESESESESESESESESESE
VEEGTPOSSALTEGDPSPALSPITELKOKELPKLPALOCRSVEEEOCNRIEET
YGVVYRAKDCKTDEIVALKRLKMEKEKGPITTSIREINTILKAOHPIVTVREIYV
SNMDKIYVMNVVEHDIKLSMETKOPPLRESEVKTIMQLRGKHLHDNWTLDK
TSTLLSHAGILKVDPECLARVSGPLKATPVVTLTRPELILGAKETSTAVDM
SYGCTFGEILTKPLPGRSEIDQINKVFKDGPSEKIMWGYSELPVKKMTSRHP
YNNLRKRGALLSDGCDLNNKFLVYFPGRRISAEDGLKHFRFTPLPIDSMFPTV
PAKSEQOVKRTSPREPGEIGYSQLODDDLKETGFLHTTNOGASAAAGPRLKF"

BASE COUNT 443 a 463 c 551 g 258 t
ORIGIN

Query Match 100.0%; Score 30; DB 9; Length 1715;
Best Local Similarity 100.0%; Pred. No. 0.3;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcgacagcgaggaagaccacg 30
|||||
Db 161 GACATCAGCGACGACGAGGAAGACCACG 190

RESULT 10
LOCUS AF067513 2226 bp mRNA PRI 07-NOV-1998
DEFINITION Homo sapiens PITSLR protein kinase alpha SV4 isoform (CDC2L1)
ACCESSION AF067513
VERSION AF067513.1 GI:3850305
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P.F. and Kidd,V.J.
TITLE Duplication of a genomic region containing the Cdc2L1-2 and
JOURNAL MRP21-22 genes on human chromosome 1p36.3 and their linkage to D122
MEDLINE Genome Res. 8 (9), 929-939 (1998)
REFERENCE 98424414
2 (bases 1 to 2226)

AUTHORS Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P.F. and Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's
Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
FEATURES
source Location/Qualifiers
1..2226
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
1..2226
/gene="CDC2L1"
620..2200
/gene="CDC2L1"
/note="Corresponding genomic sequence deposited as GenBank
Accession Numbers AF080678-AF080688, AF092429, and
AF092430"
/codon_start=1
/product="PITSLR protein kinase alpha SV4 isoform"
/protein_id="AAC72078.1"
/db_xref="GI:3850306"
/translation="MEERDLISDQIDISERKTSASSESSAEGSGSESEEEEEE
EEEGTSNSEESESEAEVESESESESESESESESESESESESESESESESE
VPESRFPDSESESEAEVEEGTPOSSALTEGDPSPALSPITELKOKELPKLPAL
OGCRSVREFQCLNRIEETGYGVVYRAKDCKTDEIVALKRLKMEKEKGPITTSIREIN
TIKAOHPIVTVREIYVSNMDKIYVMNVVEHDIKLSMETKOPPLRESEVKTIMQ
LLRGVRIHNNWILHRDLKTSNLLSHAGILKVDPECLARVSGPLKATPVVTLTR
RABPLLGAKETSTAVDMMSVGCIFGEILTKPLPGRSEIDQINKVFKDGPSEKIM
WPGYSELPAVKMTFSEHPYNNLRKRGALLSDGCDLNNKFLVYFPGRRISAEDGLK
HEYFRERPLPIDSMFPTVPAKSEQOVKRTSPREPGEIGYSQLODDDLKETGFLH
TTTNOGASAAAGPRLKF"

BASE COUNT 638 a 541 c 705 g 342 t
ORIGIN

Query Match 100.0%; Score 30; DB 9; Length 2226;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcgacagcgaggaagaccacg 30
|||||
Db 653 GACATCAGCGACGACGAGGAAGACCACG 682

RESULT 11
LOCUS AF067518 2323 bp mRNA PRI 07-NOV-1998
DEFINITION Homo sapiens PITSLR protein kinase beta SV8 isoform (CDC2L2) mRNA,
complete cds.
ACCESSION AF067518
VERSION AF067518.1 GI:3850315
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P.F. and Kidd,V.J.
TITLE Duplication of a genomic region containing the Cdc2L1-2 and
JOURNAL MRP21-22 genes on human chromosome 1p36.3 and their linkage to D122
MEDLINE Genome Res. 8 (9), 929-939 (1998)
REFERENCE 98424414
2 (bases 1 to 2323)
AUTHORS Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P.F. and Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's
Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
FEATURES
source Location/Qualifiers
1..2323
/organism="Homo sapiens"

Query Match 100.0%; Score 30; DB 6; Length 468;
Best Local Similarity 100.0%; Pred. No. 0.36;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gacatcagcgacgaggaagaccagc 30
|||||
Db 439 GACATCAGCGACGACGAGAGACACGAC 468

RESULT 6
LOCUS CDC2L2S08 480 bp DNA PRI 09-DEC-1998
DEFINITION Homo sapiens PITSURE protein kinase (CDC2L1) gene, exons 8 and 9.
ACCESSION AF080678
VERSION AF080678.1 GI:3978432
KEYWORDS
SEGMENT
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 480)
AUTHORS Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P.F. and Kidd,V.J.
TITLE Duplication of a genomic region containing the Cdc2L1-2 and
MMP21-22 genes on human chromosome 1p36.3 and their linkage to DL22
JOURNAL Genome Res. 8 (9), 929-939 (1998)
MEDLINE 98424414
REFERENCE 2 (bases 1 to 480)
AUTHORS Gururajan,R., Lahti,J.L., Grenet,J., Easton,G., Gruber,I.,
Ambros,P.F. and Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-1998) Tumor Cell Biology, St Jude Children's
Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
LOCATION/Qualifiers
FEATURES
source 1..480
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
78..173
/gene="CDC2L1"
/number=8
259..417
exon /gene="CDC2L1"
/number=9
BASE COUNT 136 a 99 c 156 g 89 t
ORIGIN

Query Match 100.0%; Score 30; DB 9; Length 480;
Best Local Similarity 100.0%; Pred. NO. 0.36;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gacatcagcgacgaggaagaccagc 30
|||||
Db 125 GACATCAGCGACGACGAGAGACACGAC 154

RESULT 7
LOCUS CDC2L2S08 592 bp DNA PRI 13-DEC-1998
DEFINITION Homo sapiens protein kinase PITSURE (CDC2L2) gene, exons 8 and 9.
ACCESSION AF080689
VERSION AF080689.1 GI:4007426
KEYWORDS
SEGMENT
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 592)
AUTHORS Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,

Ambros,P.F. and Kidd,V.J.
TITLE Duplication of a genomic region containing the Cdc2L1-2 and
MMP21-22 genes on human chromosome 1p36.3 and their linkage to DL22
JOURNAL Genome Res. 8 (9), 929-939 (1998)
MEDLINE 98424414
REFERENCE 2 (bases 1 to 592)
AUTHORS Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P.F. and Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-1998) Tumor Cell Biology, St. Jude Children's
Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
LOCATION/Qualifiers
FEATURES
source 1..592
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
78..173
/gene="CDC2L2"
/number=8
259..408
exon /gene="CDC2L2"
/number=9
BASE COUNT 154 a 126 c 184 g 128 t
ORIGIN

Query Match 100.0%; Score 30; DB 9; Length 592;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gacatcagcgacgaggaagaccagc 30
|||||
Db 125 GACATCAGCGACGACGAGAGACACGAC 154

RESULT 8
LOCUS AX033425 660 bp DNA PAT 21-SEP-2000
DEFINITION Sequence 6 from Patent WO0044896.
ACCESSION AX033425
VERSION AX033425.1 GI:10280186
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Beyaert,R. and Cornelis,S.
TITLE Internal ribosome entry site (IRES), vector containing same and
uses thereof
JOURNAL Patent: WO 0044896-A 6 03-ANG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS
SIGRID (BE)
LOCATION/Qualifiers
FEATURES
source 1..660
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 210 a 128 c 261 g 61 t
ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 660;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 gacatcagcgacgaggaagaccagc 30
|||||
Db 439 GACATCAGCGACGACGAGAGACACGAC 468

RESULT 9
AF174497

ORIGIN

BASE COUNT 11 a 8 c 10 g 1 t

Query Match 100.0%; Score 30; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.55;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagc 30
11 GACATCAGCGACGACGAGAGACCAGC 30

RESULT 2
AX033439 40 bp DNA PAT 21-SEP-2000
LOCUS AX033439 Sequence 20 from Patent WO0044896.
DEFINITION AX033439
ACCESSION AX033439
VERSION AX033439.1 GI:10280200
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequence.
REFERENCE
AUTHORS 1 (bases 1 to 40)
Beyeaert,R. and Cornelis,S.
TITLE Internal ribosome entry site (IRES), vector containing same and
uses thereof
PATENT: WO 0044896-A 20 03-AUG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS
SIGRID (BE)

FEATURES
source
Location/Qualifiers
1..40
/organism="Homo sapiens"
/db_xref="taxon:32630"
/note="5'-end primer"

BASE COUNT 14 a 10 c 12 g 4 t

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagc 30
11 GACATCAGCGACGACGAGAGACCAGC 40

RESULT 3
AX033420 222 bp DNA PAT 21-SEP-2000
LOCUS AX033420 Sequence 1 from Patent WO0044896.
DEFINITION AX033420
ACCESSION AX033420
VERSION AX033420.1 GI:10280181
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 222)
Beyeaert,R. and Cornelis,S.
TITLE Internal ribosome entry site (IRES), vector containing same and
uses thereof
PATENT: WO 0044896-A 1 03-AUG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS
SIGRID (BE)

FEATURES
source
Location/Qualifiers
1..222
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 77 a 36 c 89 g 20 t

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 222;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagc 30
1 GACATCAGCGACGACGAGAGACCAGC 30

RESULT 4
AX033421 222 bp mRNA PAT 21-SEP-2000
LOCUS AX033421 Sequence 2 from Patent WO0044896.
DEFINITION AX033421
ACCESSION AX033421
VERSION AX033421.1 GI:10280182
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 222)
Beyeaert,R. and Cornelis,S.
TITLE Internal ribosome entry site (IRES), vector containing same and
uses thereof
PATENT: WO 0044896-A 2 03-AUG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS
SIGRID (BE)

FEATURES
source
Location/Qualifiers
1..222
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 77 a 36 c 89 g 20 t

ORIGIN

Query Match 100.0%; Score 30; DB 6; Length 222;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gacatcagcagcagcagcagcagcagc 30
1 GACATCAGCGACGACGAGAGACCAGC 30

RESULT 5
AX033424 468 bp DNA PAT 21-SEP-2000
LOCUS AX033424 Sequence 5 from Patent WO0044896.
DEFINITION AX033424
ACCESSION AX033424
VERSION AX033424.1 GI:10280185
KEYWORDS
SOURCE
ORGANISM
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 468)
Beyeaert,R. and Cornelis,S.
TITLE Internal ribosome entry site (IRES), vector containing same and
uses thereof
PATENT: WO 0044896-A 5 03-AUG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS
SIGRID (BE)

FEATURES
source
Location/Qualifiers
1..468
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 144 a 100 c 182 g 42 t

ORIGIN

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:54:30 : Search time 2399.05 seconds
(without alignments)
206.296 Million cell updates/sec

Title: US-09-915-060-4

Perfect score: 30

Sequence: 1 gacatcagcagacgagaggaagaccagc 30

Scoring table: IDENTITY_NUC

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: *
1: gb_da: *
2: gb_htg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_pl: *
8: gb_pr: *
9: gb_ro: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_in: *
20: em_om: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vl: *
30: em_htgo_hum: *
31: em_htgo_inv: *
32: em_htgo_rod: *
33: em_htgo_hum: *
34: em_htg_inv: *
35: em_htg_rod: *
36: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	30	100.0	30	6	AX033423	AX033423 Sequence	
2	30	100.0	40	6	AX033439	AX033439 Sequence	
3	30	100.0	222	6	AX033420	AX033420 Sequence	
4	30	100.0	222	6	AX033421	AX033421 Sequence	
5	30	100.0	468	6	AX033424	AX033424 Sequence	
6	30	100.0	480	9	CDCL1508	AF080678 Homo sapi	
7	30	100.0	592	9	CDCL12508	AF080689 Homo sapi	
8	30	100.0	660	6	AX033425	AX033425 Sequence	
9	30	100.0	1715	9	AF174497	AF174497 Homo sapi	
10	30	100.0	2226	9	AF067513	AF067513 Homo sapi	
11	30	100.0	2333	9	AF067518	AF067518 Homo sapi	
12	30	100.0	2329	9	AF067525	AF067525 Homo sapi	
13	30	100.0	2340	9	HSU04818	U04818 Human prote	
14	30	100.0	2349	9	AF067517	AF067517 Homo sapi	
15	30	100.0	2362	9	AF067523	AF067523 Homo sapi	
16	30	100.0	2409	9	AF067521	AF067521 Homo sapi	
17	30	100.0	2436	9	AF067520	AF067520 Homo sapi	
18	30	100.0	2439	9	AF067519	AF067519 Homo sapi	
19	30	100.0	2444	9	HSU04817	U04817 Human prote	
20	30	100.0	2448	9	AF067522	AF067522 Homo sapi	
21	30	100.0	2465	9	HSU07705	U07705 Human prote	
22	30	100.0	2471	6	AX033422	AX033422 Sequence	
23	30	100.0	2471	9	HSU04816	U04816 Human prote	
24	30	100.0	2471	9	HSU07704	U07704 Human prote	
25	30	100.0	2477	9	HSU04824	U04824 Human prote	
26	30	100.0	2486	9	AF067512	AF067512 Homo sapi	
27	30	100.0	2500	9	AF067514	AF067514 Homo sapi	
28	30	100.0	2525	9	AF067515	AF067515 Homo sapi	
29	30	100.0	2533	9	AF067516	AF067516 Homo sapi	
30	30	100.0	2544	9	AK000081	AL000081 Homo sapi	
31	30	100.0	110608	9	HS283E3	AL031282 Human DNA	
32	28	93.3	35	6	AX033437	AX033437 Sequence	
33	33	26.8	89.3	1936	10	MOSP58GTA	M58633 Mouse p58/
34	34	26.8	89.3	3161	10	MUSCDPK	L37092 Mus musculi
35	35	23.6	78.7	1670	10	RATGTA	L2438 Rattus norvegicus
36	36	22.6	75.3	8071	3	MSORTIRET	M91690 Anopheles gambiae
37	37	22.6	75.3	329709	1	AP002997	AP002997 Mesorhynchus
38	38	22	73.3	248	5	CPTSLRE04	U16347 Gallus gallus
39	39	22	73.3	82852	2	AC017685	AC017685 Drosophila
40	40	22	73.3	172414	2	AC093168	AC093168 Homo sapi
41	41	22	73.3	199607	2	AC009366	AC009366 Drosophila
42	42	22	73.3	295566	3	AE003597	AE003597 Drosophila
43	43	21.8	72.7	151589	8	AC079022	AC079022 Oryza sativa
44	44	21.2	70.7	217584	1	AP002086	AP002086 Agrobacterium
45	45	21	70.0	4281	3	AY038001	AY038001 Drosophila

ALIGNMENTS

RESULT 1
AX033423
LOCUS AX033423 30 bp DNA
DEFINITION Sequence 4 from Patent WO0044896.
AX033423
ACCESSION AX033423
VERSION AX033423.1 GI:10280184
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 30)
AUTHORS Beyaert, R. and Cornelis, S.
TITLE Internal ribosome entry site (IRES), vector containing same and
uses thereof
JOURNAL VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS
SIGRID (BE)
FEATURES
source 1..30
Location/Qualifiers

Query Match	56.8%;	Score 265.8;	DB 11;	Length 826;
Best Local Similarity	85.9%;	Pred. No. 2.2e-38;		
Matches 353;	Conservative	0;	Mismatches 27;	Indels 31;
				Gaps 4

```

Query Match      56.8%; Score 265.8; DB 11; Length 826;
Best Local Similarity 85.9%; Prid. No. 2,2e-38;
Matches 353; Conservative 0; Mismatches 27; Indels 31; Caps

```

SUBMIT 12
 1259507
 CUS
 FINITION
 BE259507 674 bp mRNA EST
 601106356f1 NIH_MGC_16 Homo sapiens CDNA clone IMAGE:3342948 5
 BE259507
 BE259507.1 GI:9129943
 EST.
 human.
 ORGANISM
 Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostom
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (Passes 1 to 674)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 CDNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
 Plate: LLCM126 row: o column: 13
 High quality sequence start: 37
 High quality sequence stop: 672.

```

FEATURES
    source
        Location/Qualifiers
            1..674
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:3342948"
                /clone_lib="NIH_MGC_16"
                /tissue_type="retinoblastoma"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
                EcoRI; cDNA made by oligo-dT priming. Directionally
                cloned into EcoRI/XhoI sites using the following 5'
                adaptor: GGCACGAC(G). Library constructed by Ling Hong
                in the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: This is a NIH_MGC Library."
BASE COUNT
    248 a 115 c 229 g 82 t

```

Query Match:	43.9%:	Score 205.6:	DB 10:	Length 674:
Best Local Similarity	18.9%:	Pred. No. 1,2e-27:		
Matches 226:	Conservative 0:	Mismatches 4:	Indels 28:	Gaps 2:
QY	1	caagacatctcggaagaacgacatcggaagaacagataaagctcgcggaatggaaga	60	
Db	414	CACGAACGTCGGAAACGGCATCGAATAAACAGGATTAAGCTCCCGGAATGGAAAG	473	
QY	61	cagaagaag-aaggaatatgcgaaggaagcattccagagagaagaaggggaatgatgcygt	119	
Db	474	CAGAAAGCAAGGGAGATGCGCAAGGGAGCATACCGAGAGAA-----	516	
QY	120	gtgcctcttcaggagccgcttggagcaattgaaaggaagcgcggaagcggaagcgcaagat	179	
Db	517	-----AAGGACCGCTTGGACCAATTTGAAGGAAGCGGGAGCGGAGCGCAAGAT	566	
QY	180	gcgggagcagcagaaggaagcagtcgggaagcaaggaagcgcgagcgcggaaggaagcg	239	
Db	567	GGCGGAGCGAGCAAGAGAGACAGCGGGAGCAAAAGAGCGAGCGGCGGAGAGAGCG	626	
QY	240	gcgcgaagagcggaagcgccgcgaaggaagtgttcgcacatcacccaagc	287	
Db	627	GGCCAAAGAGCGGAGAGACCGAGGGAAGTCTTCGCATTCACCGAAC	674	

RESULT	13
LOCUS	AL034747
DEFINITION	AL034747 372 bp mRNA
ACCESSION	V8608b53 Bedlington mouse dissected endoderm Mus musculus cDNA
VERSION	clone 528_8K10 5', mRNA sequence.
KEYWORDS	AL034747
SOURCE	AL034747.1 GI:6646373
ORGANISM	EST.
	house mouse.
	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
	1 (bases 1 to 372)
TITLE	Harrison,S.M., Dunwoodie,S.L., Arkell,R.M., Lehrach,H. and
	Bedlington,R.S.
	Isolation of novel tissue-specific genes from cDNA libraries
	representing the individual tissue constituents of the gastrulating
	mouse embryo
JOURNAL	Development 121 (8), 2479-2489 (1995)
MEDLINE	95401865
COMMENT	Contact: Wiles,M., Lehrach,H. and Avner,P.
	EEC Mouse Transcript Mapping Consortium
	Genoscope - CNS
	2, rue Gaston Cremlieux, 91000 Evry, France
	Email: pavner@pasteur.fr
	clone available from Ressourcenzentrum, Heubnerweg 6, D-14059
	Berlin, Germany. Web site http://www.itzpd.de
	Seq primer: CCGGTCCGGAATTCGCGGT;

QY 61 cagaagagaaggaatgagcaaggaagcattccagagagagaaggggaatgatgagcgtg 120
 |||||
 Db 418 CAGAAGAGAAAGGAGATGCGAAGGAGACATTCCAGAGAGAA-----459
 QY 121 tgccttcaaggagccgttggagcaattcgaagaagcgggagcgaggagcaatg 180
 |||||
 Db 460 -----AGGGACCCCTTGAGCACTGTAGAAAGAGCGGAGCGGAGCCGCAAGATG 510
 QY 181 cggagagcaggaaggaagcagcggaagcagaagagcgagcgagcgaggagcgag 240
 |||||
 Db 511 CGGAGAGAGCGAAGAGAGCGCGGAGCGAGAGCGCGGAGCGCGGAGAGAGAGCGG 570
 QY 241 cgcgaagagcgaggagccgcaggaagagtgctcagacatcacggaagcgatgagagagac 300
 |||||
 Db 571 CGCAAGAGCGGAGCGCGCGAGCGCAAGTGTCTCCATCATCCGCAAGATGAGAGAGAGAC 630
 QY 301 tacagcgacaaagtgaagccagccacttgagtcgacgcccgcctcgagcgagcgaggag 360
 |||||
 Db 631 TACAGCGACAAAGTGAAAGCGAGCGACCTGGAGTGCAGCGCCGCTCGCGCGCGGAGAG 690
 QY 361 cgggttcgaattggagagcgag-cgggaagcgagtgaaagaagagaagaatggagaagaga 419
 |||||
 Db 691 CGGTTCCAGATTGGAGAGCGGTGGGAGCGAGTAAAGACAGACACACTGTCGAGCAAGG 750
 QY 420 cctgtgtccgacttaagagcatcagcgacagcgagag 458
 |||||
 Db 751 GGACCTCCGCGTCCGACTTTACAGGAGCATCAGGAGAG 789

RESULT 7
 BG107702 1340 bp mRNA EST 30-JAN-2001
 LOCUS 602217878F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4365474 5',
 DEFINITION mRNA sequence.
 ACCESSION BG107702
 VERSION BG107702.1 GI:12601548
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1340)
 AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@rsfemail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10014 row: 1 column: 19
 High quality sequence stop: 667.
 Location/Qualifiers
 1..1340
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4365474"
 /clone_lib="NIH_MGC_86"
 /tissue_type="osteocarcoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: bone; Vector: PCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.533 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

FEATURES
 source
 1..1340
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4365474"
 /clone_lib="NIH_MGC_86"
 /tissue_type="osteocarcoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: bone; Vector: PCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.533 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 BASE COUNT 467 a 272 c 450 g 151 t
 ORIGIN

Query Match 65.2%; Score 305; DB 11; Length 1340;
 Best Local Similarity 100.0%; Pred. No. 2,1e-45;
 Matches 305; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 164 agcgaggagcgcaagatgctggagcagcagaagaagcaagcgaggagcgagcgagc 223
 |||||
 Db 1 AGCGGAGCGCAAGATGCGGAGAGCAGCAGAGAGAGCGGAGAGAGAGAGCGCGAGC 60
 QY 224 ggcggcgaggaagagcgagcgaggaagcgagcgagcgaggaagtgctcagacacc 283
 |||||
 Db 61 GGGGGCGGAGAGAGCGCGCCAGAGAGCGGAGCGCGGAGAGAGTGTCTGACATCC 120
 QY 284 gaacgctgagagagagcctaacgagacaagtgaaagccagccagctggagtgccagccgc 343
 |||||
 Db 121 GAACGATGAGAGAGGACTACAGCAGCAAGTGAAGCCAGCCACTGTGATCCACCCGC 180
 QY 344 ctgcgagcgagcgagcgagcgagtgctggagagcgagcgagcaatlaaagaagaga 403
 |||||
 Db 181 CTCGGCGCGCGGAGAGCGCGTTCGAGTTGAGAGAGCGCGGAGAGCACTAAAGAGAGA 240
 QY 404 aaatggaagaaggaagcctctgtctcgcagctcaggaatcagcgacagcgagagaaga 463
 |||||
 Db 241 AAATGGAAGAAAGGAGCCTGCTGCTCGACTTACAGACATCAGCGACGAGAGAGAGA 300
 QY 464 ccagc 468
 |||||
 Db 301 CCAGC 305

RESULT 8
 AK008283 1000 bp mRNA HTC 05-JUL-2001
 LOCUS AK008283
 DEFINITION Mus musculus adult male small intestine cDNA, RIKEN full-length
 enriched library, clone:2010016J05, full insert sequence.
 ACCESSION AK008283
 VERSION AK008283.1 GI:12842370
 KEYWORDS CAP trapper.
 SOURCE Mus musculus (strain:C57BL/6J) adult male small intestine cDNA to
 mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
 clone:2010016J05.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 1000)
 AUTHORS Carninci,P. and Hayashizaki,Y.
 TITLE High-efficiency full-length cDNA cloning
 JOURNAL Methods in enzymology. 303, 19-44 (1999)
 MEDLINE 99279253
 PUBMED 10349636
 REFERENCE 2 (bases 1 to 1000)
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,K., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL Genome research. 10 (10), 1617-1630 (2000)
 MEDLINE 20499374
 PUBMED 11042159
 REFERENCE 3 (bases 1 to 1000)
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
 Kono,H., Akiyama,Y., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M.,
 Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
 Fujiwara,S., Inoue,K., Togawa,Y., Itawa,M., Ohara,E., Wataniki,M.,
 Tameda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,D.,
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format
 JOURNAL sequencing pipeline with 384 multipipillary sequencer
 MEDLINE Genome research. 10 (11), 1757-1771 (2000)
 PUBMED 20530913
 REFERENCE 4 (bases 1 to 1000)
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.

Db 721 CCGAAGGACGGGAGCCCGCAGAGCAAGTGTCTGCACATCCAGCAACGATGAGAGAGAC 780
 Qy 301 taaagcagaagatgaaacccagccactgtgagtcgacgcccctcgccgcgcgagag 360
 Db 781 TACAGCGACAAGATGAAACCCAGCCACTGAGTGCAGCCGCCCTCTCGCCGCGGAG 840
 Qy 361 cggctcaggttggaagcgcgcgaagcagctaaagaagaatgagaaga 415
 Db 841 CGGTTGAGTTGGAGACCGCCGGAAGCCAGTAAACAAAAA 895

RESULT 5
 A1580044 592 bp mRNA EST 14-DEC-1999
 LOCUS tG45908.x1 NCI-CGAP ut1 Homo sapiens cDNA clone IMAGE:2211806 3'
 DEFINITION similar to TR:Q61395 Q61399 CELL DIVISION CYCLE 2-LIKE 2 ; contains
 element TARI repetitive element ; mRNA sequence.

ACCESSION A1580044
 VERSION A1580044
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index

JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/dbp/image/image.html
 Insert length: 1082 std Error: 0.00
 Seq primer: -400p from Gibco
 High quality sequence stop: 405
 POLYA-No.

FEATURES
 source

1. 592
 /location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:2211806"
 /clone_lib="NCI-CGAP_ut1"
 /tissue_type="well-differentiated endometrial
 adenocarcinoma, 7 pooled tumors"
 /lab_host="DH10B"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: Salt;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.75 kb. Life Technologies catalog #:
 11538-014"
 58 a 212 c 115 g 207 t

Query Match 71.1%; Score 332.8; DB 10; Length 592;
 Best Local Similarity 92.8%; Pred. No. 2.3e-50;
 Matches 371; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

Qy 1 cagaacgtcggaaacacatcgaaagaacagatgaagctcgccggagatggaaga 60
 Db 373 CACGACGCTCGAAGACGCAATCGAAGAACAGATTAACCTCGCCGGATGGAGAGA 314
 Qy 61 cagaagaagaggaatcgcaagggagcattccagagagaagaaggggaaatgagcgtg 120
 Db 313 CAGAAGAGAGGAGATGGCAAGGAGCAATTCAGAGAGAGA----- 270

Qy 121 tgcctctcaggagaccgcttggagcagtltagaagaagcggagcggagcgaagatg 180
 Db 271 -----AGGACCCGCTTGAGACAGTTAGAAAGGAAGCGGAGCCGCGAGCCCAAGATG 221
 Qy 181 cggagacgcaggaagggagcagcggagcagaagagcgcgagcggcgagagcgg 240
 Db 220 CGGGAGCAGCAGAAAGGACACCGGAGCGCAAGAGCGCGAGCGGCGGAGAGACGG 161
 Qy 241 cgaagagcggagggcccgcaagggaagtgtctcacatcacccgaacgcatgagagagac 300
 Db 160 CCGAAGGAGCGGAGGAGCCCGGAGGAGTGTCTGCACATCACCGAACGATGAGAGAGAC 101
 Qy 301 tacagcgaacaagtgaagccagccactgtgagtcgcagccgctcgccgagcgag 360
 Db 100 TACAGCGACAAGTGAAGGCCAGCCACTGTGAGTGCAGCCGCCCTCGCGCGGAG 41
 Qy 361 cggctcaggttggaagcgcgcgaagcgcggaagcagctaaagaag 400
 Db 40 CGGTTGAGTTGGAGAGACGGCCGGAAGCCAGTAAAGAG 1

RESULT 6
 BG574409 804 bp mRNA EST 10-APR-2001
 LOCUS BG574409 602596352P1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4705189 5',
 DEFINITION mRNA sequence.

ACCESSION BG574409
 VERSION BG574409
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 TITLE NIH-MGC http://mgi.nci.nih.gov/
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: DCTD/DTF
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: ILAM10571 row: 9 column: 14
 High quality sequence stop: 686.

FEATURES
 source

1. 804
 /location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4705189"
 /clone_lib="NIH_MGC_87"
 /tissue_type="mammary adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SaltI; Cloned unidirectionally. Oligo-dt primed.
 Average insert size 1.383 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 272 a 165 c 270 g 97 t

Query Match 67.1%; Score 314.2; DB 11; Length 804;
 Best Local Similarity 84.5%; Pred. No. 4.9e-47;
 Matches 388; Conservative 0; Mismatches 43; Indels 28; Gaps 2;

Qy 1 cagaacgtcggaaacacatcgaaagaacagatgaagctcgccggagatggaaga 60
 Db 358 CACGACGCTCGAAGACGCAATCGAAGAACAGATTAAGCTCGCCGGATGGAGAGA 417

**JOURNAL
COMMENT**

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabds-r@mail.nih.gov
Tissue Procurement: DCTD/DPH
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: MCG Genomics, Inc.
Clone distribution: MCG clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.llnl.gov
Plate: L16K518 row: b column: 07
High quality sequence stop: 667.

BASE COUNT	371 a	159 c	390 g	122 t
ORIGIN				

Query Match	75.38;	Score 352.6;	DB 10;	Length 1042;
Best Local Similarity	90.28;	Pred. No. 6.7e-54;		
Matches 423; Conservative	0;	Mismatches 14;	Indels 32;	Gaps 3

OY		5	aacgttcggaagaaagaatcgtgaataaaaagaagtataaacctcgccgggatctggnaaaacaga	64
Dd		1	AACGTGGAAGAACGCATCGAAGAACAAGATTAAGCTGCCGGGAAITGGAAACACGA	60
OY		65	.aaagaagaagaaatgcgaag-g- gactcattccaagtagaagaagggaatgatgtgcgtgtgc	122
Dd		61	AAGAAGAGGAGTAGTGGCCAAGGTGACACTTCACGAGAGAGAA-----	99
OY		124	ccttcacagagacgcgttgagcacgttgaagaagaaccggaacggaacgcacaatatgcg	188
Dd		100	-----AAGGACCCGTTTGAGCACTTAGAAGAAAGAACCGGAGCGGAGCGCAAGTGGCG	155
OY		184	gagcagcaagaagagacagcggaagcagaagaagcgcgaacggaacggaagagacgcgc	245
Dd		154	GAGCAGCAGAAAGAGACAGCGGAGACACAGAAGACGCGGACGGCGGAGGAGCGGCC	211
OY		244	aagagcgagagcgcccgcgaaggaaatgtctcgcacatccagaaacyaataagagaactac	303
Dd		214	AAGGAGCGGGAGGCCCCGCAAGGAAAGTCTGCACATCACCGAACAATGAGAGAGACTAC	273
OY		304	agcgacaagaatgaaagccagaccacttgtatcgtgaagccgcgcccgccgcgcgggaaggg	365
Dd		274	AGCGACAAAGTGAAGCCAGCCACTGGAGTGCAGCCGCCCTCTGGCGCGCGGGAGGG	333
OY		364	ttcgaatctggagaacgcgcggaagccqta----aagaagaanaatlgyaagaagaagga	419
Dd		334	TTCGAGTTGGAGAGCGGCCGGAAGCCAGTAAACGAAGAGAACACTGGCAAGACAAAGGA	399
OY		420	cctgtgtctcgacttaagaacatccagcgaacagcgaagagaaagaccagc	468
Dd		394	CCTGCTGTCTCGACTTAAGAAGCATACCGCAGCACCGAGAGAAAGACCAC	442

RESULT	4
BC009375	
LOCUS	
BC009375	906 bp mRNA
DEFINITION	Homo sapiens, similar to cell division cycle 2-like 1 (PITSLRE
	proteins), clone IMAGE:4121554, mRNA.

ACCESSION
VERSION
KEYWORDS
SOURCE

KEYWORDS
SOURCE

ORGANISM

REFERENCE
Eukaryota; Metazoa: Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria: Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 906)

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 906)
Strausberg, R.
Direct Submission
Submitted (12-JUN-2001) National Institutes of Health, Mammalian

REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk

CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: <http://www.nisc.nih.gov/>
Contact:

Shevchenko, Y., Welchby, K.D., Beststrom-Stenberg, S.M., Benjamin, B., Blakesley, R.M., Bouffard, G.G., Brinkley, C., Brooks, S. Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Kallins, E., Legaspi, R. Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stancirip, S., Thomas, P. J., Tlousong, E.E., Touchman, J.W., Tsungeoan, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 26 Row: d Column: 15
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 3650303
 This clone has the following problem: Incomplete processing.

FEATURES	Location/Qualifiers
source	1. .906

BASE COUNT	317 a	170 c	307 g	112 t
ORIGIN				

Query Match	71.9%;	Score 336.6;	DB 12;	Length 906;
Best Local Similarity	91.3%;	Pred. No. 4.8e-51;		
Matches 379; Conservative	0;	Mismatches 9;	Indels 27;	Gaps 1;

QY	1	cacgaagcctcggaaacgcgcatctcgagaagaaacagatataaagctcgcgcggaaatgggaaga	60
Db	508	CACGAACGTGCGAAMCGCATCTCGAGAAGAACAGATATAAGCTCCCGGAA TGCAAGA	567
QY	61	cagaagagaaggaaattgcaagaggacattccagagagagaagaaagggagatgatgycgt	120
Db	568	CAGAAGAGAGGAGATTGCAGAGGAGACATTCCAGGAGAGAA-----	607
QY	121	tgcctcttcagagaccgcttgtagcacgtatagaagaagacgsgagcggagagcgcaagaatg	180
Db	610	-----AGGGACCGCTTGAGCAGTGTGAAGAAAGGAAACGGGAGCGGAGCGCAAGTC	660
QY	181	cggagacagcagaagagacagcggagacagaagaagacgacagcggcggagcgagagagcgg	240
Db	661	CGGAGACAGCAEAAAGAGACACGGGAGACAAGAGACCGGACCGGCGCGAGAGACGG	720
QY	241	cgcagaagacgagagcccgaagaaqtcttgcacatccacccaacatatagaagaagac	300
Db			

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:14:19 ; Search time 3881.49 seconds
(without alignments)
1295.643 Million cell updates/sec

Title: US-09-915-060-5

Perfect score: 468

Sequence: 1 cacgacgctcggaacgaca.....acagcgagaggaagacacgc 468

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estlom:*
5: em_estlpl:*
6: em_estda:*
7: em_estro:*
8: em_estov:*
9: em_hlrc:*
10: qd_estcl:*
11: qd_estc2:*
12: qd_hlrc:*
13: qd_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	396.8	84.6	1063	10	BE740559 601595656
2	395.8	84.6	723	10	BE742222 601575629
3	352.6	75.3	1042	10	BE742943 601574925
4	336.6	71.9	906	12	BC009375 Homo sapi
5	332.8	71.1	592	10	AI580044 t445908.x
6	314.2	67.1	804	11	BG574409 602596352
7	305	65.2	1340	11	BG107702 602277878
8	295.2	63.1	1000	12	AK008283 Mus muscu
9	281	60.0	662	11	BG822824 602727816
10	271.6	58.0	326	11	BG943585 ax39f04.x
11	265.8	56.8	826	11	BG034678 602300028
12	205.6	43.9	674	10	BE259507 601106356

13	203	43.4	372	10	AL034747	AL034747 v8608b53
14	200.4	42.8	812	11	BI149024	BI149024 602911290
15	186.6	39.9	186	6	AA681955	AA681955 v444b10.s
16	172.8	36.9	516	11	BF156008	BF156008 RC0-HT095
17	171	36.5	830	11	BF607177	BF607177 MT1-00011
18	162	34.6	758	11	BG969359	BG969359 602836841
19	155.6	33.2	1083	11	BG770237	BG770237 602744821
20	155.2	33.2	493	11	BI018825	BI018825 IL3-MT026
21	153	32.7	547	11	BF743770	BF743770 PM0-BT085
22	138	29.5	441	10	AK451883	AK451883 UT-H-B13-
23	127.6	27.3	725	11	BF308197	BF308197 601887445
24	124	26.5	514	10	AI989489	AI989489 ws25d04.x
25	120.8	25.8	462	11	BG486162	BG486162 dsd23e11.
26	120.8	25.8	476	10	BE680738	BE680738 d184d07.y
27	119.6	25.6	251	10	AV270235	AV270235 AV270235
28	119.4	25.5	732	10	AI006623	AI006623 uc16a10.y
29	110.2	23.5	160	10	AM193827	AM193827 xm30g12.x
30	106	22.6	673	11	BF983533	BF983533 602306504
31	104.4	22.3	469	11	BF359805	BF359805 RC6-MT006
32	103.8	22.2	321	10	AA309711	AA309711 EST180769
33	102.4	21.9	233	10	BB290670	BB290670 BB290670
34	100.4	21.5	253	10	AV327502	AV327502 AV327502
35	96.4	20.6	372	11	BF918032	BF918032 IL3-UT011
36	95	20.3	292	10	AV368285	AV368285 AV368285
37	86.8	18.5	479	11	BF934664	BF934664 IL2-NP020
38	85.2	18.2	253	10	AV331446	AV331446 AV331446
39	85	18.2	182	10	AA13247	AA13247 v52d12.r
40	83.8	17.9	702	10	BE308367	BE308367 601090048
41	80.4	17.2	202	10	AV330095	AV330095 AV330095
42	76.4	16.3	987	13	CNS00418	AL066537 Dicosoph11
43	74.4	15.9	997	13	CNS0057E	AL067067 Dicosoph11
44	73	15.6	740	10	AL043340	AL043340 DRFP4340
45	71.8	15.3	912	13	CNS02LEG	AL202705 Tetradon

ALIGNMENTS

RESULT 1
BE740559 1063 bp mRNA EST 15-SEP-2000
LOCUS BE740559 601595656f1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3949693 5',
DEFINITION BE740559
ACCESSION BE740559
VERSION BE740559.1 GI:10154551
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1063)
NIH-MGC <http://mgc.ncl.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DMR
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
Plate: LNC814 row: p column: 14
High quality sequence stop: 741.
Location/Qualifiers
1..1063
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3949693"
/clone_lhb="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:

FEATURES

source

Db 2911 GGCAGAGCAGGAGGGCGCAGAGAGGTGGAGCCCGGGGTCTGAGGAGGCGAGTGGAGG 2970
QY 320 ccag 323
Db 2971 CCGG 2974

RESULT 10
US-08-194-087-15
; Sequence 15, Application US/08194087
; Patent No. 5879910
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/194,087
; FILING DATE: 18-MAY-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 779
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-194-087-15

Query Match 12.3%; Score 57.6; DB 2; Length 10596;
Best Local Similarity 49.3%; Pred. No. 0.00014;
Matches 150; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 20 atcgaagaacagataaagctcgccggaatgggaagacagaagaaggaatgg 79
Db 2671 ACCAGAGAGGCGCAGAGAGGAGGCGCAGAGAGGCGGCGAGGAGGCGGCGAGG 2730
QY 80 caagggagatcccaaggaggaaggggaatgatgctgtgtgctcttcagggaacgct 139
Db 2731 AGGGGAGAGAGGCGCAGAGAGGAGGCGGCGAGGAGGCGCAGAGAGGAGGCGGCGG 2790
QY 140 tggagcagttagaagaagcaggaagcaggaatgcaggaagcagaagaagagc 199
Db 2791 AGGGGAGAGAGGCGCAGAGAGGAGGCGGCGCAGAGAGGAGGCGGCGAGGCGGCGG 2850
QY 200 agcgggaagcagaagagcaggaagcaggaagcaggaagcaggaagcaggaagc 259
Db 2851 ACCAGAGAGGCGCAGAGAGGAGGCGGCGCAGAGAGGCGGCGAGGAGGCGGCGAGG 2910
QY 260 gcaagggaagtgctctgcaatcaaccgaacgatgaggaaggaactaacgcagcaaatgaaag 319

Db 2911 GGCAGAGCAGGAGGGCGCAGAGAGGTGGAGCCCGGGGTCTGAGGAGGCGAGTGGAGG 2970
QY 320 ccag 323
Db 2971 CCGG 2974

RESULT 11
PCT-US93-04648-15
; Sequence 15, Application PC/TUS9304648
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc., Godowski, Paul J., Lokker, Nathalie A., Mark, Melani
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/04648
; FILING DATE: 19930517
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/884811
; FILING DATE: 18-MAY-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/885971
; FILING DATE: 18-MAY-92
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 755,779P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10596 bases
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US93-04648-15

Query Match 12.3%; Score 57.6; DB 5; Length 10596;
Best Local Similarity 49.3%; Pred. No. 0.00014;
Matches 150; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 20 atcgaagaacagataaagctcgccggaatgggaagacagaagaaggaatgg 79
Db 2671 AGCGAGAGGCGCAGAGAGGAGGCGGCGCAGAGAGGAGGCGGCGAGGAGGCGGCGAGG 2730
QY 80 caaaggagcatlcccaaggagaagaaggggaatgatgctgtgtgctcttcagggaacgct 139
Db 2731 AGGGGAGAGAGGCGCAGAGAGGAGGCGGCGCAGAGAGGAGGCGGCGAGGAGGCGGCGG 2790
QY 140 tggagcagttagaagaagcaggaagcaggaatgcaggaagcagaagaagagc 199
Db 2791 AGGGGAGAGAGGCGCAGAGAGGAGGCGGCGCAGAGAGGAGGCGGCGAGGAGGCGGCGG 2850
QY 200 agcgggaagcagaagagcaggaagcaggaagcaggaagcaggaagcaggaagc 259
Db 2851 AGCAGAGAGGCGCAGAGAGGAGGCGGCGCAGAGAGGAGGCGGCGAGGAGGCGGCGAGG 2910

QY	320	ccag	323
Db	2971	CCGG	2974

```

? RESULT      8
? US-08-087-783A-15
? Sequence 15, Application US/08087783A
? Patent No. 554/856
? GENERAL INFORMATION:
? APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
? TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR VARIANTS
? NUMBER OF SEQUENCES: 22
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Genentech, Inc.
? STREET: 460 Point San Bruno Blvd
? CITY: South San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94080
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: winpatin (genentech)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/087,783A
? FILING DATE: 13-Jul-1993
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/884811
? FILING DATE: 18-MAY-92
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/885971
? FILING DATE: 18-MAY-92
? ATTORNEY/AGENT INFORMATION:
? NAME: Marschang, Diane L.
? REGISTRATION NUMBER: 35,600
? REFERENCE/DOCKET NUMBER: P0755779P1
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415/225-5416
? TELEFAX: 415/952-9881
? TELEX: 910/371-7168
? INFORMATION FOR SEQ. ID NO.: 15:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 10596 base pairs
? TYPE: Nucleic Acid
? STRANDEDNESS: Single
? TOPOLOGY: Linear
? US-08-087-783A-15

```

	Query Match	12.3%	Score 57.6	DB 1	Length 10596	
	Best Local Similarity	49.3%	Prod. No. 0.00014			
	Matches 150	Conservative	0	Mismatches 154	Indels 0	Gaps 0
QY	20	atcgaagaacaggatataagctcgccggaatgggaagaacagaagaaggaatg	79			
Db	2671	AGCAGAGAGGGGACGAGCAGGAGGGGACAGACAGAGGGGCCAGGAGGGGACGAGCAG	2730			
QY	80	caaggaagcattccagagagagaaggggaatgatggtgtgtctcttcagggaacct	139			
Db	2711	AGGGCAGAGAGGGGACGAGCAGGAGGGGCCAAGAGGACGAGAGGAGGGGACAG	2790			
QY	140	tggagcaattgaagaagaagcggagcgcggaagcgaagatgcyggagcagcgaagaagc	199			
Db	2791	AGGGCAGAGAGGACGAGGAGGAGGGGACGAGGAGGGGCCAAGAGGAGGGGCCAGG	2850			
QY	200	agcgggaagcagaagaagcggcgagcgcgcggaagagcggcgcaagaagaagcgagcc	259			
Db	2851	AGCAGAGAGGGGACGAGGAGGGGACGAGCAGGAGGGGCCAAGAGGAGGAGGAGG	2910			
QY	260	gaaggaaggtgtctcaatccagcaagcatgaaggaagcctacagcgcaagaagtgaag	319			

Accession	Sequence	Position
Db	2911 GGACGAGCAGAGGGCCAGAGAGAGTGGAGGCCGGGTCGAGAGCACCACGAGAG	2970
Q7	320 ccag 323	
Db	2971 CCGG 2974	

```

1      RESULT          9
2      US-08-194-088B-15
3      ; Sequence 15, Application US/08194088B
4      ; Patent No. 5580963
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Godowski, Paul J. Lokker, Nathalie A. Mark, Melanie R.
7      ; TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS
8      ; NUMBER OF SEQUENCES: 21
9      ; CORRESPONDENCE ADDRESS:
10     ; ADDRESSEE: Genentech, Inc.
11     ; STREET: 460 Point San Bruno Blvd
12     ; CITY: South San Francisco
13     ; STATE: California
14     ; COUNTRY: USA
15     ; ZIP: 94080
16
17     ; COMPUTER READABLE FORM:
18     ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
19     ; COMPUTER: IBM PC compatible
20     ; OPERATING SYSTEM: PC-DOS/MS-DOS
21     ; SOFTWARE: patin (Genentech)
22     ; CURRENT APPLICATION DATA:
23     ; APPLICATION NUMBER: US/08/194, 088B
24     ; FILING DATE: 09-FEB-1994
25     ; CLASSIFICATION: 530
26
27     ; PRIOR APPLICATION DATA:
28     ; APPLICATION NUMBER: 07/884811
29     ; FILING DATE: 18-MAY-1992
30     ; ATTORNEY/AGENT INFORMATION:
31     ; NAME: Gallegos, R. Thomas
32     ; REGISTRATION NUMBER: 32,692
33     ; REFERENCE/DOCKET NUMBER: 755D1
34     ; TELECOMMUNICATION INFORMATION:
35     ; TELEPHONE: 415/225-2614
36     ; TELEFAX: 415/952-9881
37     ; TELEX: 910/371-7168
38     ; INFORMATION FOR SEQ ID NO: 15:
39     ; SEQUENCE CHARACTERISTICS:
40     ; LENGTH: 10596 bases
41     ; TYPE: nucleic acid
42     ; STRANDEDNESS: single
43     ; TOPOLOGY: linear
44     ; US-08-194-088B-15

```

Query Match	12.3%	Score 57.6	DB 1	Length 10596
Best Local Similarity	49.3%	Prod. No. 0.00014		
Matches 150	Conservative	Mismatches 154	Indels	Gaps
QY	20	atcgcgaagaacagaagatataagctcgcgcgggaatggaagaacagagaagaaggaatg	79	
Db	2671	ACCAGAGAGGGCGAGAGCAGGAGGGGCGAGACAGAGAGGGGCGAGGCGCAGGACGAG	2730	
QY	80	caaggaagacatccacgcagagagaagaaggggaatgatgctgtgtcccttcagggaacgct	139	
Db	2731	AGGGGCGAGAGGGGCGAGAGCGAGGAGGGGCGAGGAGGCGACAGCAGGAGAGGGGCGAG	2790	
QY	140	tggagcagcttgaagaagaagcggagcgcggaagcgcaaatgcggaagcagacagaagag	199	
Db	2791	AGGGCGAGGACGAGGAGGAGGGGCGAGGAGGGGCGACAGCAGGAGGGCGCAGGAGGGGCGAG	2850	
QY	200	agcgggaagcagaagaagcgcgaagcggcgcggaaggaagcgcgcaagaagcgggaagccc	259	
Db	2851	ACCAGAGAGGGCGAGGAGGGGCGAGGAGGAGGGGCGACAGGAGGGCGCAGGAGGAGG	2910	
QY	260	gcagggaagtgctcacatcaccaacagatbagaaggaagactacaagcgacaacaagtgaag	319	

[illegible]

```

RESULT      4
US-09-130-114-1/c
: Sequence 1, Application US/09130114
: Patent No. 5976807
: GENERAL INFORMATION:
: APPLICANT: Horlick, Robert A.
: APPLICANT: Dama, Bassam B.
: APPLICANT: Robbins, Alan K.
: TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
: TITLE OF INVENTION: From Multiple Transfected Episomes
: FILE REFERENCE: 0867/1D903US1
: CURRENT APPLICATION NUMBER: US/09/130,114
: CURRENT FILING DATE: 1998-08-06
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 3452
: TYPE: DNA
: ORGANISM: EBVNA
US-09-130-114-1

```

	Query Match	12.3%	Score 57.6:	DB 2:	Length 5452:
	Best Local Similarity	49.3%:	Pred. NO.0.0013:		
	Matches 150:	Conservative 0:	Mismatches 154:	Indels 0:	Gaps 0:
QY	20 atcgaagaacacagatataagctccgcggaattggaaagacaagaagaaaggaatlg	79			
Db	1714 AGCAGGAGGGCGACGACGAGGAGGGGCCAGACGACGAGGCGCAGGAGGCAGACGAG	1655			
QY	80 caaggagcatlccagsgaagaagaagggaatlgtatgtgcgttcltcaaggaccgct	139			
Db	1654 AGGGCCAGGAGGGGGCGAGGACGAGGGGGSCAAGAGGGGGCACGAGCAGGAGGAGGGCCAGG	1595			
QY	140 tggagcatlttgaaagagcggggacggggcggggacgatgtgggaagcacgaagaygc	199			
Db	1594 AGGGCCAGGACGACGAGGAGGGGGCGAGAGGGCCAGCAGCAAGAGGGGGCGAGGGGGCCAGG	1535			
QY	200 agcgggaaccaaagagcgcgcgcgcggaagacgycgcaagaagcgyagggccc	259			
Db	1534 AGCAGGAGGGCGCAGGAGGGGGCGAGGACGAGAGGGGGCAGGAGCGACGAGAGGAG	1475			
QY	260 gcaaggaaagtctctgcacatcacccaagcatgaagagactacagcgaaatgaag	319			

Dd	1474	GGCAGGACGAGAGGGCGACGAGAGTGGAGGCCGGGTGAGGAGCGCTGTGAGG	1415
QY	320	ccag 323	
Dd	1414	CCGG 1411	

```

RESULT      5
US-08-910-647-1
: Sequence 1, Application US/08910647
: Patent No. 6251433
: GENERAL INFORMATION:
:   APPLICANT: Zuckermann et al.
:   TITLE OF INVENTION: Compositions and Methods for
:   TITLE OF INVENTION: Polynucleotide Delivery
:   NUMBER OF SEQUENCES: 4
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: Chiron Corporation
:     STREET: 4560 Horton Street
:     CITY: Emeryville
:     STATE: California
:     COUNTRY: U.S.A.
:     ZIP: 94608-2916
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: Floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC-DOS/MS-DOS
:     SOFTWARE: PatentIn Release #1.0, Version #1.30
:   CURRENT APPLICATION DATA:
:     APPLICATION NUMBER: US/08/910,647
:     FILING DATE:
:     CLASSIFICATION: 514
:   ATTORNEY/AGENT INFORMATION:
:     NAME: Fujita, Sharon M.
:     REGISTRATION NUMBER: 38,459
:     REFERENCE/DOCKET NUMBER: 1218.002
:   TELECOMMUNICATION INFORMATION:
:     TELEPHONE: (510) 923-2706
:     TELEFAX: (510) 655-3542
:   INFORMATION FOR SEQ. ID NO.: 1:
:     SEQUENCE CHARACTERISTICS:
:       LENGTH: 9600 base pairs
:       TYPE: nucleic acid
:       STRANDEDNESS: single
:       TOPOLOGY: linear
:     MOLECULE TYPE: DNA (genomic)
US-08-910-647-1

```

Query Match	12.3%	Score 57.6	DB 4	Length 9600
Best Local Similarity	49.3%	Pred. No. 0.00014		
Matches 150	Conservative 0	Mismatches 134	Indels 0	Gaps 0
QY	20	atcgaagaacacagataaagctccgcgcggaatgynaagacagaagaagaaatcgtg	79	
Db	1137	AGCAGAGAGGGCCAGGAGCAGGAGGGGCGAGAGCAGAGAGGGCCAGAGGACGAGG	1196	
QY	80	caaggaagcattccagaagaagaaggaatgtagtgcgtgtgcctcttcagaagccgt	139	
Db	1197	AGGGGCGAGAGGGGCGAGGACGAGGCGGCGAGAGGGGCGACGAGAGGAGGCGGCGAGG	1256	
QY	140	tgaagcagttcgaagaaggaagcgcggaagcgcgaatgctcgggaagcagcagaagaagc	199	
Db	1257	AGGGCGAGAGCAGAGAGGAGGGGCGAGGAGGGGCGAGGACGAGAGGAGGCGGCGAGG	1316	
QY	200	agcgcggaagcagaagaagagcgcgagcgcgcggaagagcgcgcagaagagcgcggaagccc	259	
Db	1317	AGCAGAGAGGGGCGAGAGGCGGCGAGGAGCGAGGAGGGGCGAGGAGGAGGAGGAGG	1376	
QY	260	gcagaggaagtgctgcacatcacccgaagaatgagagaaggaactacagcgcacaagtgaag	319	
Db	1377	GCGAGAGACAGAGGGGCGAGGAGCAGAGAGAGTGAGAGCCGGGGGTGAGAGAGCGATGGAGG	1436	

```
Query Match 21.2%; Score 99.4; DB 1; Length 7218;
Best Local Similarity 5.4%; Pred. No. 7.3e-14;
Matches 22; Conservative 259; Mismatches 130; Indels 0; Gaps 0;

Qy 2 acgaagctcggaacgacatcggaagaacagataagctcgccggaatggaagac 61
Db 1469 ATGCACTACTTTAAAGATAGAGATTTCGACRRRRRRRRRRRRRRRRRRRR
Qy 62 agaagagaaggaatgcaaggaagcattccagagagaaggaaggaatgagcgtg 121
Db 1409 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
Qy 122 gctcttcggggcgccgtggagcagltgaaaggaagcgggagcggaagatgc 181
Db 1349 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
Qy 182 ggaagacagaaggaagcagcggaagcagaagagcgcgagcggaagcgcgc 241
Db 1289 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
Qy 242 gcaagagcgggagcgccgcaaggagtgctgcacatccagacgatgagagag 301
Db 1229 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
Qy 302 acagcgacaaagtgaagccagccactgagtcgacagccgcctcgccgcggaag 361
Db 1169 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
Qy 362 ggttcgagtgaggagcgccgcaagcagtaagaagaagaatggaag 412
Db 1109 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
RESULT 2
US-08-458-568A-11/c
; Sequence 11, Application US/08458568A
; Patent No. 5821339
; GENERAL INFORMATION:
; APPLICANT: Schaffer, Priscilla A.
; TITLE OF INVENTION: Compositions and Methods for Treatment of Herpesvirus
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5821339rits
; STREET: One liberty place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,568A
; FILING DATE: 02-JUNE-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/065,146
; FILING DATE: 05-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn R.
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: DFCI-0029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12001 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Herpes simplex virus
; STRAIN: Herpes Simplex Virus Type 1
; US-08-458-568A-11
```

```
Query Match 12.4%; Score 58; DB 1; Length 12001;
Best Local Similarity 57.9%; Pred. No. 0.00012;
Matches 103; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
```

```
Qy 79 gcaagggagcattccagagagaaggaatgagcgtgtgctcttcagggaccgc 138
Db 1541 GCGGGCGCGCGCGCGCAAAAAGCGCGCGCGCGCGCGCGCGCGCGCGCG
Qy 139 ttgagcagttagaaggaagcgaggagcgcaagatgagcgagcagaagag 198
Db 1481 CGGGCGTGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGAG 1422
Qy 199 cagcggaagcagaagagcgagcgagcgagcgagcgagcgagcgagcgag 256
Db 1421 GAGCGGGGAGAGAGAGCGGGGAGAGAGAGCGGGGAGAGAGCGGGGAGAG 1364
```

```
RESULT 3
US-09-050-863-2
; Sequence 2, Application US/09050863
; Patent No. 6114111
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hiang, Betty
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,863
; FILING DATE: 30-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
; US-09-050-863-2
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:57:33 ; Search time 152.61 Seconds
(without alignments)
694,527 Million cell updates/sec

Title: US-09-915-060-5
Perfect score: 468
Sequence: 1 caccagacgtcggaacgaca.....acagcgagaggaagaccagc 468

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/prodata/2/1na/5A_COMB.seq:*
2: /cgn2_6/prodata/2/1na/5B_COMB.seq:*
3: /cgn2_6/prodata/2/1na/6A_COMB.seq:*
4: /cgn2_6/prodata/2/1na/6B_COMB.seq:*
5: /cgn2_6/prodata/2/1na/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/2/1na/backfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	99.4	21.2	7218	1 US-08-232-463-14	Sequence 14, Appl
C 2	58	12.4	12001	1 US-08-458-568A-11	Sequence 11, Appl
C 3	57.6	12.3	2580	3 US-09-050-863-2	Sequence 2, Appl
C 4	57.6	12.3	5452	2 US-09-130-114-1	Sequence 1, Appl
C 5	57.6	12.3	9600	4 US-08-910-647-1	Sequence 1, Appl
C 6	57.6	12.3	10596	1 US-07-884-811-15	Sequence 15, Appl
C 7	57.6	12.3	10596	1 US-07-885-971-15	Sequence 15, Appl
C 8	57.6	12.3	10596	1 US-08-087-783A-15	Sequence 15, Appl
C 9	57.6	12.3	10596	1 US-08-194-088B-15	Sequence 15, Appl
C 10	57.6	12.3	10596	2 US-08-194-088B-15	Sequence 15, Appl
C 11	57.2	12.2	10596	5 PCT-US93-04648-15	Sequence 15, Appl
C 12	57.2	12.2	9551	2 US-08-056-200-93	Sequence 93, Appl
C 13	57.2	12.2	9551	2 US-08-800-644-93	Sequence 93, Appl
C 14	55.6	11.9	16442	3 US-08-781-891-208	Sequence 208, App
C 15	54.8	11.7	397	3 US-09-253-691-3	Sequence 3, Appl
C 16	54	11.5	51259	3 US-08-781-891-209	Sequence 209, App
C 17	53.8	11.5	2301	5 US-08-306-691B-23	Sequence 23, Appl
C 18	53.8	11.5	2301	5 PCT-US93-08251-78	Sequence 78, Appl
C 19	53	11.3	289	4 US-09-007-005-17	Sequence 17, Appl
C 20	53	11.3	289	4 US-09-244-796-17	Sequence 17, Appl
C 21	51.4	11.0	168	1 US-08-469-802B-4	Sequence 4, Appl
C 22	51.4	11.0	168	2 US-08-267-803B-4	Sequence 4, Appl
C 23	51.4	11.0	171	1 US-08-469-802B-5	Sequence 5, Appl
C 24	51.4	11.0	171	2 US-08-267-803B-5	Sequence 5, Appl
C 25	51.4	11.0	195	1 US-08-469-802B-2	Sequence 2, Appl
C 26	51.4	11.0	195	2 US-08-267-803B-2	Sequence 2, Appl
C 27	51.4	11.0	234	1 US-08-469-802B-3	Sequence 3, Appl

28	51.4	11.0	234	2 US-08-267-803B-3	Sequence 3, Appl
29	51	10.9	3376	1 US-08-320-559-29	Sequence 29, Appl
30	51	10.9	3376	3 US-08-545-860D-29	Sequence 29, Appl
31	51	10.9	3376	5 PCT-US94-04496-29	Sequence 29, Appl
32	50	10.7	203	4 US-09-043-303-7	Sequence 1, Appl
33	49.4	10.6	477	4 US-09-135-994-1	Sequence 1, Appl
34	49.2	10.5	3489	2 US-08-728-323A-1	Sequence 1, Appl
35	49.2	10.5	32207	2 US-08-770-379-20	Sequence 20, Appl
C 36	49.2	10.5	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 37	48.4	10.3	154	1 US-08-469-802B-6	Sequence 6, Appl
C 38	48.4	10.3	154	2 US-08-267-803B-6	Sequence 6, Appl
C 39	48.2	10.3	1931	2 US-09-130-114-2	Sequence 2, Appl
C 40	47.8	10.2	165	4 US-09-043-303-17	Sequence 17, Appl
C 41	45.4	9.7	543	6 5273901-6	Sequence 1, Appl
C 42	44.8	9.6	53526	3 US-08-658-136-2	Sequence 2, Appl
C 43	44.8	9.6	53577	3 US-08-658-136-1	Sequence 1, Appl
C 44	44.6	9.5	250	1 US-08-480-552-16	Sequence 16, Appl
C 45	44.6	9.5	250	3 US-08-929-208-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)883-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpt-FLS
US-08-232-463-14

DR WPI: 1995-022825/03.
XX
XX Herpes Simplex Virus (HSV) specific junction spanning transcript
PT - for inhibiting HSV L/ST synthesis, in the treatment of HSV
PT Infection.
XX
XX
PS Disclosure; Page 38-44; 64pp; English.
XX
XX An HSV-specific junction-spanning transcript (L/ST) maps at the 5'
CC end to the b repeat sequences of HSV DNA at approx. 3 and 125 kb,
CC and extends into the c repeat sequences of HSV DNA. Compounds
CC that inhibit HSV L/ST synthesis may be used as anti-HSV
CC virucides. The HSV-1 DNA sequence in the region of the L/STs
CC is given in AA076213.
XX
XX
SQ Sequence 12001 BP; 1568 A; 4557 C; 4315 G; 1561 T; 0 other:

Query Match 12.4%; Score 58; DB 16; Length 12001;
Best Local Similarity 57.9%; Pred. No. 0.00021;
Matches 103; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 79 gcaaggaagcattccaggaagaaggaatgagtgctgtgctcttcagggaccgc 138
DB 1541 GCGGGCGCGCGCGCGCAAAAAGCGGGCGGCGTCCGGCGCGCTGCGCGCGCGCG 1482
QY 139 ttggaagcagttagaagaaggaagcaggaagtcggaagcagcagaagga 198
DB 1481 CGGGCGTGGGGGGCGGGCGCGGAGCGGGGAGAGAGCGGGGGGAGCGGGGGAG 1422
QY 199 cagcgaggaagcaggaagcaggaagcaggaagcaggaagcaggaagga 256
DB 1421 GAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAGAGCGGGGAGAGCGGGGGG 1364

RESULT 14
AAD07132
ID AAD07132 standard; cDNA; 2803 BP.
XX
XX
AC AAD07132;
XX
DT 06-AUG-2001 (first entry)
XX
DE Canine retinitis pigmentosa GTPase regulator (RPGR) mutant #2 cDNA.
XX
KW Dog: X-linked progressive retinal atrophy 2; XLPR2; genetic marker;
KW retinitis pigmentosa GTPase regulator; RPGR; Siberian Husky; Samoyed;
KW Miniature Schnauzer; mutant; muteln; ss.
XX
OS Canis familiaris.
OS Synthetic.
XX
FH Key Location/Qualifiers
CDS 1..1149
FT /*tag- a
FT /*product= "Canine retinitis pigmentosa GTPase
FT regulator mutant"
FT /*note- "CDS does not include start codon"
FT /partial
FT replace (931..932, AGAG)
FT /*tag- b
FT /*note- "This deletion results in the change of
FT amino acids and ends in a premature stop codon"
XX
XX
PN MO200138578-A1.
XX
XX
PD 31-MAY-2001.
XX
PF 21-NOV-2000; 2000MO-US31940.
XX
PR 24-NOV-1999; 99US-0167365.
XX
XX (CORR) CORNELL RES FOUND INC.

XX
PI Aguirre GD, Acland GM, Zhang Q, Ray K, Zeiss CJ;
XX
XX WPI: 2001-367707/38.
DR
DR P-PSDB; AAE02399.
XX
XX
PT Identifying dogs with or carrying X-linked progressive retinal atrophy
PT by detecting retinitis pigmentosa GTPase regulator gene mutation,
PT useful when breeding Husky, Samoyed and Miniature Schnauzer -
PS Claim 75; Page 34-35; 88pp; English.
XX
XX
XX The invention relates to a method for identifying dogs which are
CC genetically normal, are carriers of, or are affected with X-linked
CC progressive retinal atrophy (XLPR), by testing a biological sample with
CC genetic markers that co-segregate with a XLPR gene locus. The invention
CC also relates to canine retinitis pigmentosa GTPase regulator (RPGR), RPGR
CC mutants and their corresponding nucleic acid molecules. The mutated RPGR
CC genes are responsible for the XLPR in dogs. Methods are used to select
CC dogs for breeding so that dogs carrying the mutated locus are eliminated
CC from the breeding stock. The method particularly applies to Siberian
CC Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
CC the disease is X-linked. XLPR1 type is identified in Siberian Huskies,
CC and Samoyeds, while XLPR2 type is identified in Miniature Schnauzers.
CC The present cDNA sequence is the mutant open reading frame (ORF) 15
CC encoding Canine retinitis pigmentosa GTPase regulator (RPGR) mutant
CC found in XLPR2-affected dogs. This mutant is obtained by deleting
CC 'GA' nucleotide bases from position 932 to 933 of the wild-type
XX canine RPGR cDNA.
XX
SQ Sequence 2803 BP; 1028 A; 287 C; 939 G; 549 T; 0 other:

Query Match 12.4%; Score 57.8; DB 22; Length 2803;
Best Local Similarity 51.8%; Pred. No. 0.00017;
Matches 131; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 47 ggaatgggaagacagacagaagaaggaatgcaagggagcattccaggaagaagga 106
DB 913 gaggaaaggaagaatgaaatgaaaggaagaagaagaagcaggggaaggggaagga 972
QY 107 ggaatgagtgctgtgctcttcacagaccgcttgagcagtlagaagaagcggagc 166
DB 973 ggaaggaaggggggaagcagggcaggaagaaggggaaggaaggggaaggaagga 1032
QY 167 gggagggcagaagtgccgggagcaggaagaagcagcgggaaggaaggaagcgagc 226
DB 1033 agggaggggaagggggcgaaggaaggaagaagaaggggaaggaagaaggaagga 1092
QY 227 gggcaggaagcagcgccgaaggaagcggcaggaagtgctcacatcacgaa 286
DB 1093 gggggggggaaggggaaggaaggggaaggaaggaaggaaggaagggaggaagga 1152
QY 287 cgatgagaagga 299
DB 1153 ggaagaggaagta 1165

RESULT 15
AAV55831
ID AAV55831 standard; DNA; 799 BP.
XX
XX
AC AAV55831;
XX
DT 18-NOV-1998 (first entry)
XX
XX
DE Nucleotide sequence of the stabilising sequence-encoding insert.
XX
KW Fusion protein: stabilising polypeptide; proteolytic degradation;
KW resistance: half-life; autoimmune disease; inflammation; nitro drug;
KW IkappaB regulator protein; inflammatory bowel disease; in vivo imaging;
KW nitroreductase protein; enzyme therapy; prodrug therapy; protease;
KW cancer; pathological condition; ss.

Db	441	gcgcgcgcgaag	450	
RESULT 10				
AA12156	1D	AAA12156 standard; DNA; 2875 BP.		
XX	XX	AAA12156;		
XX	XX	10-AUG-2000 (first entry)		
XX	XX	Human X5L DNA fragment.		
XX	XX	X5L; human; spermatogenesis protein; treatment; diagnosis; detection; ds.		
XX	XX	Homo sapiens.		
XX	XX	DE19856882-CL1.		
XX	XX	04-MAY-2000.		
XX	XX	10-DEC-1998; 98DE-1056682.		
XX	XX	10-DEC-1998; 98DE-1056682.		
XX	XX	(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.		
XX	XX	Sedlacek Z, Poustka A;		
XX	XX	WPI: 2000-294180/26.		
XX	XX	P-PSDB; AAA12156.		
XX	XX	New spermatogenesis protein useful for investigation and modulation of		
XX	XX	spermatogenesis, particularly for diagnosis and treatment of		
XX	XX	spermatogenesis disorders -		
XX	XX	Claim 4; Fig 2; 14pp; German.		
XX	XX	This invention describes a novel spermatogenesis protein (I) and the DNA		
XX	XX	molecule (II) encoding (I). (I) and (II) are used to investigate and		
XX	XX	modulate spermatogenesis, especially to activate or inhibit it and for		
XX	XX	diagnosis and treatment of spermatogenesis disorders. (I) is also used		
XX	XX	to raise specific antibodies, useful for detection of (I) and as		
XX	XX	therapeutic inhibitors of (I) and to detect specific autoantibodies. (II)		
XX	XX	can be used as a source of primers to determine the organization and		
XX	XX	expression of the (I)-encoding gene, also for in vivo expression of (I).		
XX	XX	This sequence encodes the human spermatogenesis protein X5L which is		
XX	XX	described in the method of the invention.		
SQ		Sequence 2875 BP; 616 A; 797 C; 894 G; 568 T; 0 other;		
Query Match 12.7%; Score 59.6; DB 21; Length 2875;				
Best Local Similarity 66.2%; Pred. No. 6.6e-05;				
Matches 86; Conservative 0; Mismatches 44; Indels 0; Gaps 0;				
QY	117	cggtgtgccttcaaggagccgttggagcagcttaagaaggaagcgggagcggaagcga 176		
Db	1359	cctgaagcaatgaaggccggcaggaagcccttggtcagggaagcgcgagcgagctggc 1418		
QY	177	gattcgaggaacacaaagaagcagcaggaagcagaagaagcgcgagcgcgggcggaaga 236		
Db	1419	caagcgagcagcactctggaagaacgcgctcagcaggaagcgcgagcaggaagcagaaga 1478		
QY	237	gcgcgcgcaag 246		
Db	1479	gcgcgcgcgaag 1488		
RESULT 11				
AA007131	1D	AA007131 standard; cDNA; 2800 BP.		
XX	XX			

AC AAD07131;
XX
DT 06-AUG-2001 (first entry)
XX
DE Canine retinitis pigmentosa GTPase regulator (RPGR) mutant #1 cDNA.
XX
KM Dog; X-linked progressive retinal atrophy 1; XLRPA1; genetic marker;
retinitis pigmentosa GTPase regulator; RPGR; Siberian Husky; Samoyed;
XX Miniature Schnauzer; mutant; muteln; ss.
OS
XX Canis familiaris.
OS Synthetic.FH
FH Key Location/Qualifiers
FT 1..882 {
FT CDS {
FT /*tag= a
FT /product= "Canine retinitis pigmentosa GTPase
FT regulator mutant"
FT /note= "CDS does not include start codon"
FT /partial
FT mutation replace (877..878, AGGAGT)
FT /*tag= b
FT /note= "This deletion replaces Arg with Ile and results
in a premature stop codon"FT
FT
FT
FT
FT
FT
FT
PD WO200138578-AI.
XX
XX 31-MAY-2001.
XX
PF 21-NOV-2000; 2000WO-US31940.
XX
PR 24-NOV-1999; 99US-0167365.
PA (CORR) CORNELL RES FOUND INC.
PI Aguirre GD, Acland GM, Zhang Q, Ray K, Zeiss CJ;
DR WPI; 2001-367707/38.
XX P-PSDB; AAEO2398.
XX
PT Identifying dogs with or carrying X-linked progressive retinal atrophy
PT by detecting retinitis pigmentosa GTPase regulator gene mutation,
PT useful when breeding Husky, Samoyed and Miniature Schnauzer
PS Claim 73; Page 32-33; 88pp; English.
XX
XX The invention relates to a method for identifying dogs which are
CC genetically normal, are carriers of, or are affected with X-linked
CC progressive retinal atrophy (XLRPA), by testing a biological sample with
CC genetic markers that co-segregate with a XLRPA gene locus. The invention
CC also relates to canine retinitis pigmentosa GTPase regulator (RPGR), RPGR
CC mutants and their corresponding nucleic acid molecules. The mutated RPGR
CC genes are responsible for the XLRPA in dogs. Methods are used to select
CC from the breeding stock so that dogs carrying the mutated locus are eliminated
CC from the breeding stock. The method particularly applies to Siberian
CC Husky, Samoyed and Miniature Schnauzer breeds, or any other breed where
CC the disease is X-linked. XLRPA1 type is identified in Siberian Huskies,
CC and Samoyeds, while XLRPA2 type is identified in Miniature Schnauzers.
CC The present cDNA sequence is the mutant open reading frame (ORF) 15
CC encoding Canine retinitis pigmentosa GTPase regulator (RPGR) mutant
CC found in XLRPA1-affected dogs. This mutant is obtained by deleting
CC 'AGGAA' nucleotide bases from position 878 to 882 of the wild-type
CC canine RPGR cDNA.
XX
XX Sequence 2800 BP; 1026 A; 287 C; 938 G; 549 T; 0 other;

Query Match 12.7%; Score 59.4; DB 22; Length 2800;
Best Local Similarity 52.2%; Pred. No. 7.3e-05;
Matches 132; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

47 gggaatgggaaagacagaagaagaagggaatgcgaaggcattccaggagagaaggg 106

CC used for treating autoimmune diseases, cancer and inflammation. In
CC particular, the core protein may be an Ikappab regulator protein for the
CC treatment of inflammatory bowel disease, or a nitroreductase protein
CC which can activate nitro drugs in enzyme/prodrug therapy to treat cancer
CC or other pathological conditions. The fusion proteins can also be used in
CC diagnostic methods such as in vivo imaging.

SO Sequence 795 BP; 200 A; 104 C; 478 G; 13 T; 0 other;

Query Match 13.2%; Score 61.8; DB 19; Length 795;
Best Local Similarity 48.7%; Pred. No. 1.6e-05;
Matches 168; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

QY 25 gaagaacaggataaagctcgcgggaatggaaagacagagaagaagaatggcaag 84
DB 429 gaaggcagcagagggcagagcagagggcagagcagagcagagggcagagggg 488
QY 85 gagcattccagcagagaagaagggaatgatgctgtctcttcacggacgcgttgag 144
DB 489 caggagcagagaggcagagggcagagcagagcagagggcagagggcagagcagag 548
QY 145 cagttagaagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 204
DB 549 ggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 608
QY 205 gagcagaagaagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 264
DB 609 gagggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 668
QY 265 gaagtgctcacatcacccgaacgatgagaggaactacacgcaaaagtgaagccagc 324
DB 669 caggagagggcagagaggcagagcagagagagaggcagagggcagagcagagcagag 728
QY 325 cactggagtcgacccctccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 369
DB 729 aggttgagggcggcggtcgcagagcagtgagggcggcggtcgcgcgcgcgcgcgcgc 773

RESULT 5

AAAF22443
ID AAFF22443 standard; cDNA; 968 BP.

XX AAF22443;

XX 26-MAR-2001 (first entry)

XX Human breast cancer associated antigen nucleotide sequence SEQ ID NO:22.

XX Human: breast cancer; gastric cancer; prostate cancer; diagnosis;

XX KW cancer associated antigen; cytostatic; cancer vaccine; ss.

XX OS Homo sapiens.

XX WO200073801-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US14749.

XX 28-MAY-1999; 99US-0136526.

XX PR 10-SEP-1999; 99US-0153454.

XX PA (LUDW-) LUDWIG INST CANCER RES.

XX Obata Y;

XX WPI; 2001-025274/03.

PT Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer -

XX Claim 50; Page 278; 799pp; English.

CC AAFF22422 to AAFF22626, AAFF22627 to AAFF22773 and AAFF22774 to AAFF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer.

SO Sequence 968 BP; 219 A; 289 C; 329 G; 127 T; 4 other;

Query Match 12.7%; Score 59.6; DB 22; Length 968;
Best Local Similarity 66.2%; Pred. No. 5.3e-05;
Matches 86; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 117 cgtgtgctcttcacggaccgcttgagcagttagaagaagcggagcggagcgca 176
DB 164 ccttgacgacacatgaaggccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 223
QY 177 gatcgaggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 236
DB 224 caagcgccgacccctggaggaagcagcggctgcagcagagcgcgcgcgcgcgcgcgc 283
QY 237 gcgcgcgcag 246
DB 284 gcgcgcgcag 293

RESULT 6

AAAI2155
ID AAAI2155 standard; cDNA; 1618 BP.

XX AAAI2155;

XX 10-AUG-2000 (first entry)

XX Human X5L cDNA.

XX X5L; human; spermatogenesis protein; treatment; diagnosis; detection; ss.

XX OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 113..1090

XX FT /*tag= a

XX FT /product= "spermatogenesis protein X5L"

XX DE19856882-C1.

XX 04-MAY-2000.

XX 10-DEC-1998; 98DE-1056882.

XX 10-DEC-1998; 98DE-1056882.

XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Sedlacek Z, Poustka A;

XX WPI; 2000-294180/26.

XX P-PSDB; AAY87646.

PT New spermatogenesis protein useful for investigation and modulation of
PT spermatogenesis, particularly for diagnosis and treatment of
PT spermatogenesis disorders -
XX Claim 3a; Fig 1; 14pp; German.

```
XX AC AAN3711;
XX 07-DEC-2000 (first entry)
XX DE Human PITSURE protein kinase isoform alpha2-2.
XX KW IRES; internal ribosome entry site; PITSURE; human; gene therapy;
XX cancer; restenosis; p58; p110; protein kinase; ds.
XX OS Homo sapiens.
XX FH key Location/Qualifiers
XX FT misc_signal 907..1128
XX FT /*tag= a
XX FT /label= IRES
XX FT /note= "internal ribosome entry site"
XX PN WO200044896-A1.
XX PD 03-AUG-2000.
XX PE 26-JAN-2000; 2000WO-EP00643.
XX PR 26-JAN-1999; 99EP-0200216.
XX PA (VLAAR-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX PI Cornelis S, Beyaert R;
XX DR WPI; 2000-49931/44.
XX PT Nucleic acids encoding internal ribosome entry sequences useful for
XX PT directing protein expression in gene therapy procedures -
XX PS Disclosure; Page 32-33; 57pp; English.
XX CC The PITSURE protein kinase gene can be translated to give two isoforms,
XX CC p110 and p58. Transcription of p58 mRNA was found to start at an
XX CC internal ribosome entry site (IRES). The IRES element was found to be
XX CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX CC cycle. The present sequence contains the human PITSURE protein kinase isoform
XX CC alpha2-2. This sequence contains the IRES. The IRES sequence and
XX CC fragments of it may be used to induce cell cycle dependent initiation
XX CC of translation in eukaryotic cells. Vectors containing the IRES may be
XX CC used for the preparation of compositions for the treatment of and/or
XX CC prevention of a disease by gene therapy. Such diseases may be cancer
XX CC and restenosis.
XX SO Sequence 2471 BP; 697 A; 601 C; 809 G; 364 T; 0 other;
Query Match 100.0%; Score 468; DB 21; Length 2471;
Best Local Similarity 100.0%; Pred. No. 2,5e-99;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 cagcaagctcgaaacgacatcgagaagaacagataaagctcgcggaatgggaaga 60
DB 469 cagcaacgctcggaacgacatcgagaagaacagataaagctcgcggaatgggaaga 528
QY 61 cagaagaagaagaatgcaagaaggagcattccaggagagaagaagggaatgagcgtg 120
DB 529 cagaagaagaagaatgcaagaaggagcattccaggagagaagaagggaatgagcgtg 588
QY 121 tgcctctcaggagcgccttgagcagttagaagaaggagcgaggcggaagcgaagatg 180
DB 589 tgcctctcaggagcgccttgagcagttagaagaaggagcgaggcggaagcgaagatg 648
QY 181 cggagagcagcagaagcagcagcagcagaagaagcgcgagcgcgagcggaagcgag 240
DB 649 cggagagcagcagaagcagcagcagcagaagaagcgcgagcgcgagcggaagcgag 708
QY 241 cgcaagagcggaagcccgagcaggaagtgctgcacatccacgcaacgatgagaagagc 300
```

```
DB 709 cgcaagagcggaagcccgcaagggaagtgtctgcacatccgcaacgatgagaagagc 768
QY 301 tacagcgacaagaatgaaagccagccactgagtgcgacccgcgctcgccgcgagag 360
DB 769 tacagcgacaagaatgaaagccagccactgagtgcgacccgcgctcgccgcgagag 828
QY 361 cggctcgagctgggagcagcgccggaagccagttaaaagaagaatggaagaagagac 420
DB 829 cggctcgagctgggagcagcgccggaagccagttaaaagaagaatggaagaagagac 888
QY 421 ctgcttcgcgacttacagacatcagcagcagcagcagaaggaagaccagc 468
DB 889 ctgcttcgcgacttacagacatcagcagcagcagcagaaggaagaccagc 936
RESULT 4
AAV55830
ID AAV55830 standard; DNA; 795 BP.
XX AC AAV55830;
XX DT 18-NOV-1998 (first entry)
XX DE FLGA insert stabilising polypeptide encoding DNA.
XX KW Fusion protein; stabilising polypeptide; proteolytic degradation;
XX KW resistance; half-life; autoimmune disease; inflammation; nitro drug;
XX KW IkappaB regulator protein; inflammatory bowel disease; in vivo imaging;
XX KW nitroreductase protein; enzyme therapy; prodrug therapy; protease;
XX KW cancer; pathological condition; ss.
XX OS Epstein-barr virus.
XX FH key Location/Qualifiers
XX FT CDS 1..788
XX FT /*tag= a
XX FT /product= "stabilising polypeptide"
XX PN MO9822577-A1.
XX PD 28-MAY-1998.
XX PE 17-NOV-1997; 97WO-1B01508.
XX PR 25-JUN-1997; 97US-0048945.
XX PR 15-NOV-1996; 96US-0030986.
XX PA (MASU/) MASUCCI M G.
XX PI Masucci MG;
XX DR WPI: 1998-312463/27.
XX DR P-FSDB; AAM79128.
XX PT New fusion proteins resistant to proteolytic degradation -
XX PT comprising a core protein with a stabilising polypeptide comprising
XX PT a peptide sequence containing glycine repeats
XX PS Disclosure; Fig 3; 120pp; English.
XX CC This DNA encodes a stabilising polypeptide and is the FLGA insert of the
XX CC invention. The invention provides a method for increasing the resistance
XX CC of a core protein to proteolytic degradation that comprises linking or
XX CC inserting onto or into the core protein a stabilising polypeptide of
XX CC formula ((Gly)X(Gly)Y(Gly)Z)n where Glya, Glyb, Glyc are 1-6
XX CC sequential Gly residues and X, Y, Z are Ala, Ser, Val, Ile, Leu, Met,
XX CC Phe, Pro or Thr and n can be anything between 1-66. X, Y and Z need not
XX CC be identical from n repeat to n repeat. Alternatively a nucleic acid
XX CC encoding the stabilising polypeptide can be linked onto or inserted into
XX CC a nucleic acid encoding a core protein. The fusion proteins of the
XX CC invention are more resistant to degradation by proteases and, thus, have
XX CC a longer half-life than the unfused core protein. The products can be
```


CC The PITSURE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence contains a 5' fragment of the PITSURE
CC protein kinase IRES and nucleotides upstream of the IRES. The IRES
CC sequence and fragments of it may be used to induce cell cycle
CC dependent initiation of translation in eukaryotic cells. Vectors
CC containing the IRES may be used for the preparation of compositions for
CC the treatment of and/or prevention of a disease by gene therapy. Such
CC diseases may be cancer and restenosis.

XX
XX
SQ Sequence 468 BP; 144 A; 100 C; 182 G; 42 T; 0 other;

Query Match 100.0%; Score 468; DB 21; Length 468;
Best Local Similarity 100.0%; Pred. No. 1,8e-99;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caggaacgtcgggaacacatcgagaagaacagataaagctgcgggaatggaaga 60
DB 1 caggaacgtcgggaacacatcgagaagaacagataaagctgcgggaatggaaga 60
QY 61 cagaagaaggaggaatgcaaggagcattccaggagaagaaggaggaatgagcgtg 120
DB 61 cagaagaaggaggaatgcaaggagcattccaggagaagaaggaggaatgagcgtg 120
QY 121 tgcctcttcaggagaccgtctgagacgttagaagaacgagagcgagagcgcaagatg 180
DB 121 tgcctcttcaggagaccgtctgagacgttagaagaacgagagcgagagcgcaagatg 180
QY 181 cgggagagcagaagaagcagcagcgagcagaagaagcgcgagcgcgagagagcg 240
DB 181 cgggagagcagaagaagcagcagcgagcagaagaagcgcgagcgcgagagagcg 240
QY 241 cgaagaagcagcagagcccgcaaggagtgctgcacatcacccagcagtggaagagac 300
DB 241 cgaagaagcagcagagcccgcaaggagtgctgcacatcacccagcagtggaagagac 300
QY 301 taacagcagacaagaatgaaagccagccactgagtgagtcgagccgcctcgcgcgcgagag 360
DB 301 taacagcagacaagaatgaaagccagccactgagtgagtcgagccgcctcgcgcgcgagag 360
QY 351 cggctcgaagttgggagacggccggaagccagttaaagaagaataatggaagaagagac 420
DB 351 cggctcgaagttgggagacggccggaagccagttaaagaagaataatggaagaagagac 420
QY 421 ctgctgtccgacttaacagacatcagcgacagcgagaagaaagaccagc 468
DB 421 ctgctgtccgacttaacagacatcagcgacagcgagaagaaagaccagc 468

RESULT 2
AAA73714
ID AAA73714 standard; DNA; 660 BP.

AC AAA73714;

DT 07-DEC-2000 (first entry)

DE Human PITSURE protein kinase gene IRES and upstream sequence.

XX IRES; internal ribosome entry site; PITSURE; human; gene therapy;

KW cancer; restenosis; p58; p110; protein kinase; ds.

XX Homo sapiens.

OS

FT misc_signal

Location/Qualifiers

439..660

/tag= a

/label= IRES

/note= "internal ribosome entry site"

FT
FT
FT
XX

PN WO200044896-A1.
XX
XX 03-AUG-2000.
XX
XX 26-JAN-2000; 2000WO-EP00643.
XX
XX 26-JAN-1999; 99EP-0200216.
XX
XX (VLA-) VLAAMS INTERNUNIVERSITAIR INST BIOTECHNOG.
XX
XX Cornells S, Beyaert R;
XX
XX WPI; 2000-499331/44.
XX
XX

PT Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -
PS Claim 10; Page 34; 57pp; English.

XX The PITSURE protein kinase gene can be translated to give two isoforms,
XX p110 and p58. Transcription of p58 mRNA was found to start at an
XX internal ribosome entry site (IRES). The IRES element was found to be
XX cell cycle regulated, p58 is produced during the G2/M stage of the cell
XX cycle. The present sequence contains the PITSURE protein kinase
XX IRES and upstream nucleotides. The IRES sequence and fragments of it
XX may be used to induce cell cycle dependent initiation of translation in
XX eukaryotic cells. Vectors containing the IRES may be used for the
XX preparation of compositions for the treatment of and/or prevention of
XX a disease by gene therapy. Such diseases may be cancer and restenosis.

SQ Sequence 660 BP; 210 A; 128 C; 261 G; 61 T; 0 other;

Query Match 100.0%; Score 468; DB 21; Length 660;
Best Local Similarity 100.0%; Pred. No. 1,9e-99;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caggaacgtcgggaacacatcgagaagaacagataaagctgcgggaatggaaga 60
DB 1 caggaacgtcgggaacacatcgagaagaacagataaagctgcgggaatggaaga 60
QY 61 cagaagaaggaggaatgcaaggagcattccaggagaagaaggaggaatgagcgtg 120
DB 61 cagaagaaggaggaatgcaaggagcattccaggagaagaaggaggaatgagcgtg 120
QY 121 tgcctcttcaggagaccgtctgagacgttagaagaacgagagcgagagcgcaagatg 180
DB 121 tgcctcttcaggagaccgtctgagacgttagaagaacgagagcgagagcgcaagatg 180
QY 181 cgggagcagcagaagaagcagcagcgagcagaagaagcgcgagcgcgagagagcg 240
DB 181 cgggagcagcagaagaagcagcagcgagcagaagaagcgcgagcgcgagagagcg 240
QY 241 cgaagaagcagcagagcccgcaaggagtgctgcacatcacccagcagtggaagagac 300
DB 241 cgaagaagcagcagagcccgcaaggagtgctgcacatcacccagcagtggaagagac 300
QY 301 taacagcagacaagaatgaaagccagccactgagtgagtcgagccgcctcgcgcgcgagag 360
DB 301 taacagcagacaagaatgaaagccagccactgagtgagtcgagccgcctcgcgcgcgagag 360
QY 351 cggctcgaagttgggagacggccggaagccagttaaagaagaataatggaagaagagac 420
DB 351 cggctcgaagttgggagacggccggaagccagttaaagaagaataatggaagaagagac 420
QY 421 ctgctgtccgacttaacagacatcagcgacagcgagaagaaagaccagc 468
DB 421 ctgctgtccgacttaacagacatcagcgacagcgagaagaaagaccagc 468

RESULT 3
AAA73711
ID AAA73711 standard; DNA; 2471 BP.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 19:03:36 : Search time 355.05 Seconds
(without alignments)
1130.061 Million cell updates/sec

Title: US-09-915-060-5

Perfect score: 468

Sequence: 1 cagcagctgcggaacgaca.....acagcagaggaagaccagc 468

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_1101:*

1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	468	100.0	468	21	AAAT3713
2	468	100.0	660	21	AAAT3714
3	468	100.0	2471	19	AAAT3711
4	61.8	13.2	795	19	AAV55830
5	59.6	12.7	968	22	AAAT2443
6	59.6	12.7	1618	21	AAAT2155
7	59.6	12.7	1638	21	AAAT6636
8	59.6	12.7	1645	22	AAAT58935
9	59.6	12.7	1648	22	AAAT60701
10	59.6	12.7	2875	21	AAAT2156
11	59.4	12.7	2800	22	AAAT07131

12	59.4	12.7	2805	22	AAAT07130	Canine retinitis p
13	58	12.4	12001	16	AAAT07213	HSV L/ST region.
14	57.8	12.4	2803	22	AAAT07132	Canine retinitis p
15	57.6	12.3	799	19	AAAT55831	Nucleotide sequenc
16	57.6	12.3	1926	21	AAAT50254	Epstein Barr virus
17	57.6	12.3	1926	22	AAAT82902	EBV tethering prot
18	57.6	12.3	2580	21	AAAT5454	Nucleotide sequenc
19	57.6	12.3	5452	20	AAAT09923	Anti-sense strand
20	57.6	12.3	8705	20	AAAT23778	Vector pshuttle DN
21	57.6	12.3	9600	19	AAAT21683	Vector plasmid pcm
22	57.6	12.3	10380	20	AAAT22248	Nucleotide sequenc
23	57.6	12.3	10596	14	AAAT01731	Plasmid pCISBON f
24	57.6	12.3	10596	17	AAAT40348	Plasmid pCISBON f
25	57.6	12.3	10596	20	AAAT15650	Nucleotide sequenc
26	57.6	12.3	16080	21	AAAT59553	DNA clone pCEK C1.
27	57.2	12.2	9551	20	AAAT22301	CDNA encoding a hu
28	56.2	12.0	1925	20	AAAT30924	Epstein Barr Virus
29	56	12.0	699	21	AAAT36119	Partial mouse WRN
30	55.6	11.9	16442	18	AAAT83006	Human c-erbB oncogene
31	54	11.5	51259	18	AAAT83007	Human Trk oncogene
32	53.8	11.5	2301	19	AAAT04445	Arabidopsis thalia
33	53.8	11.5	2301	21	AAAT09303	Spinocherebellar at
34	53.4	11.4	696	21	AAAT48939	Murine LOBO homolo
35	53.2	11.4	397	20	AAAT89891	Spinocherebellar at
36	51.6	11.0	49999	20	AAAT23895	Spinocherebellar at
37	51.4	11.0	171	16	AAAT84833	Spinocherebellar at
38	51.4	11.0	195	16	AAAT84831	Spinocherebellar at
39	51.4	11.0	234	16	AAAT84832	Gene encoding a su
40	51.2	10.9	10732	21	AAAT10594	Probe #15454 for g
41	51	10.9	465	22	AAAT25521	Probe #2079 used
42	51	10.9	465	22	AAAT15993	Probe #6274 for ge
43	51	10.9	486	22	AAAT16341	Probe #7609 used t
44	51	10.9	486	22	AAAT16341	
45	51	10.9	486	22	AAAT38923	

ALIGNMENTS

RESULT 1

AAAT3713

ID AAAT3713 standard; DNA; 468 BP.

XX

AC AAAT3713;

XX

DT 07-DEC-2000 (first entry)

XX

DE 5' extension of PITSURE protein kinase internal ribosome entry site.

XX

KW IRES; internal ribosome entry site; PITSURE; human; gene therapy;

KM cancer; retinoblastoma; p53; p110; protein kinase; ds.

XX

OS Homo sapiens.

XX

XX WO2000044896-A1.

PN

PD 03-AUG-2000.

XX

XX 26-JAN-2000; 2000MO-EP00643.

XX

XX 26-JAN-1999; 99EP-0200216.

PR

XX (VIAA-) VIAAMS INTERNUNIVERSITAIR INST BIOTECHNOG.

PA

XX Cornelis S, Beyaert R;

PI

XX WPI; 2000-499331/44.

DR

XX Nucleic acids encoding internal ribosome entry sequences useful for

XX directing protein expression in gene therapy procedures -

PS Claim 9; Page 33; 57pp; English.

Wed Dec 5 08:45:08 2001

us-09-915-060-5.rge

Page 11

Db 909 CTGCTGTCCGACTTACAGACATCAGCGACGACGAGAGAGAACACCAGC 956

Search completed: December 4, 2001, 18:54:37
Job time: 9173 sec

	Db	829	CAGTTCGAGTTGGAGACGCCCGGAAGCCAGCTAAAAAGAAGAAAATGAAAGAAAGGC	888
Oy	421	ctcgtctccgacttacagacatcacgcagcagcgagaagaaccagc	468	
Db	889	CTGCTGTCCGACTTACAGACATCATCAGCCACACGAGAGAGAAAGCACAGC	936	
RESULT	13			
LOCUS	HSU04824	2477 bp	mRNA	PRI 08-JUL-1994
DEFINITION	Human protein kinase PITSLRE alpha 2-1 mRNA, complete cds.			
ACCESSION	U04824			
VERSION	U04824.1	GI:507167		
KEYWORDS	human.			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 2477)			
AUTHORS	Xiang,J., Lahiri,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.			
TITLE	Molecular cloning and expression of alternatively spliced PITSLRE protein kinase isoforms			
JOURNAL	J. Biol. Chem. 269, 15786-15794 (1994)			
MEDLINE	94253170			
REFERENCE	2 (bases 1 to 2477)			
AUTHORS	Kidd,V.J.			
TITLE	Direct Submission			
JOURNAL	Submitted (04-JAN-1994) Vincent J. Kidd, St. Jude Children's Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St., Memphis, TN 38101, USA			
FEATURES	location/Qualifiers			
SOURCE	1..2477			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/map="lp36-2"			
	/cell_line="HeLa S3"			
	/cell_type="epithelioid"			
	/tissue_type="cervix"			
	/clone_lib="HeLa cDNA library"			
	/dev_stage="adult"			
	1..2477			
	112..2451			
	/codon_start=1			
	/product="PITSLRE alpha 2-1"			
	/protein_id="AA19586.1"			
	/db_xref="gi:507168"			
	/translation="MGDEKDSWKYKTLDLLEIQEKRRKEQEKAIKKLKNSDDRS*			
	RDSLEEGELRHDCMEITIRNPSRREDSEMERGEDSDLAIKPOOMSKRKVHIHKI			
	EKKREKRHRSHSGAGCKHARVKKERHERKRRERODCARMEORCKRWMAREHE			
	SRPRDLLEQLERRERPKMRBOOKRQRERARRAEPERKREARRRVSAHHRTM			
	REDYSKYKASHMSRSPRPFRERELGDGRKPYEEMERDLSLDIDISDRKTK			
	SSAASSSSGSSEGESEEEEEEEESTSESSESESEEEETGSSSEASOSAS			
	EVSEESSEDERENENHLIVPPSRFSDSSEAEVEEGEPQSSALTEGVYVA			
	DSPALSPELKEOELPKYLALQCRSVSEFCILNRIEGTGYVARADKTDIELVAL			
	KRLMEKKEGFPTSLREITNLIAKHNPNTAVYARETVVGSNMKIYVMNVYEHDIK			
	SLMTFMKOPFLPGEVKTLIMLDLGSKVLHNWMLIHRLDKSNLLISAGLIKYDFEE			
	LARETSPLAKITFYVVTLMTRAPLELGAKETADVDMVSVCIFGLLTQKPLFPK			
	KSELIDQINKYFKRIDGPSKTIWPYSELPAVKMTFSHPNNLRKRGALLSDGET			
	LNMFLETFPGRRISAEDELKHEYFRFPLPIDTSMPTTWPAKSBOOKVYKRGTSR			
	EGGIGYSOLIGDDCLKETGFIHTTNOCGSAAGPGFSIKF"			
BASE COUNT	701 a 603 c 811 g 362 t			
ORIGIN				
Query Match	86.3%; Score 404; DB 9; Length 2477;			
Best Local Similarity	94.2%; Pred. No.2.le-69;			
Matches	441; Conservative 0; Mismatches 0; Indels 27; Gaps 1;			
Oy	1	cacgaagtcggnaacgaacatcagaaagacaagataaacgtccgcgggaatcggnaaga	60	
Db	502	CACGAACCTCGGAACGACATCGAAGAACGAGGTCTTAAGCTCGCCGGGAATGGGAAGA	561	

QY	61	cagaagaagaaggaatgtggcaaggaagcattcccggaagaagaagggggatcgaagcgtg	120
Db	562	CAGAAGACAAGGGAATGTGCAAGGGAGCATTTCCAGAGAGAA-----	603
QY	121	tgcccttcaggaaccgcttgagcagcttaagaagaagcgagcgagcgacgaatg <th>180</th>	180
Db	604	-----AGGAGCCGCTTGAGCAGATTAGAAAGAACGGGACCGGACCCGACGATG	654
QY	161	cggagacagcaagaagaacgacgaggaagcaagaagacgagcgagcgagcgaggaagcg <th>240</th>	240
Db	655	CGGAGCAGCAGAAAGACACGCGGGAGCAGAAGGAGCCGAGCGGCGGGAGAGCGG <th>714</th>	714
QY	241	cgcaagagcgagggagcccgcaaggaatgtctgacacatccgaaagatagaagagac <th>300</th>	300
Db	715	CGCAAGGAGCGGGAGCCCGCAGGGAATGTCTGCATCATCCGAAACATATGAGAGAGAC <th>774</th>	774
QY	301	tacaacgcaagaagttaaaagccagccacatcgaggtgcgaaccccgctcgagcgcgagag <th>360</th>	360
Db	775	TACACGCAAAAGTAAAGCCAGCCACTGAGATGCGACCCCGCTCGCCCGCGGAG <th>834</th>	834
QY	361	cggctcagatctgggagagcgccgaggaacgaatgaagaagaagaatgtgaagaagagac <th>420</th>	420
Db	835	CGGTTCCGAGTTGGGAGAGCGCCGCGAGACGCAGTAAAGAAAGAAATGAGAAAGGAC <th>894</th>	894
QY	421	ctgcgtccgacattacagacatcaagcagcagcgaggaagaccagc <th>468</th>	468
Db	895	CTGCTGTCGAGCTTACAGGACATCATGCCACGACGAGAGAGAACCCAGC <th>942</th>	942
RESULT	14		
LOCUS	AF067512		
DEFINITION	Homo sapiens PIRSLRE protein kinase alpha SV1 isoform (CDC2L1)		
ACCESSION	AF067512		
VERSION	AF067512.1		
KEYWORDS	GI:3850303		
ORGANISM	human.		
SOURCE	Homo sapiens		
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.		
AUTHORS	1 (bases 1 to 2486)		
TITLE	Gururajan, R., Lahli, J. M., Grenet, J., Easton, J., Gruber, I., Ambros, P. F., and Kidd, V. J.		
JOURNAL	Duplication of a genomic region containing the Cdc2L1-2 and MM21-22 genes on human chromosome 1p36.3 and their linkage to DL22		
MEDLINE	Genome Res. 8 (9), 929-939 (1998)		
FEATURES	98424414		
SOURCE	2 (bases 1 to 2486)		
TITLE	Gururajan, R., Lahli, J. M., Grenet, J., Easton, J., Gruber, I., Ambros, P., and Kidd, V. J.		
JOURNAL	Submitted		
MEDLINE	Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA		
AUTHORS	Location/Qualifiers		
TITLE	1. 2486		
JOURNAL	/organism="Homo sapiens"		
MEDLINE	/db_xref="taxon:9606"		
FEATURES	/chromosome="1"		
SOURCE	/map="1p36.3"		
gene	1. 2486		
CDS	/gene="CDC2L1"		
	112.. 2460		
	/gene="CDC2L1"		
	/note="Corresponding genomic sequence deposited as GenBank		
	Accession Numbers AF080678-AF080688, AF092429, and		
	AF092430"		
	/codon_start=1		
	/product="PIRSLRE protein kinase alpha SV1 isoform"		
	/translation="MGEDEKSSVYKTLIDELIQRKRKEQEEKAEIKRLKNSDNRD		
	/db_xref="GI:3850304"		
	/protein_id="AAC72077.1"		
	/RDSIEBEELDHMEITIRNSPYRAREMDRDEEDSLAIKPPQOMSREKVVHNRKD		

QY	1	cacgaagcttcggaacacgacatcgcgaagaagaacaggaataagctccgcggaaatggaaga	60
Db	464	CACGAAGCTCGGAACGACATCGAAGAAACAGGATTAAGCTCCCGGAAATGGGAAGA	523
QY	61	cagaagagaaggaatgcaagaagagagatctccagagagagaagaaggggaatgatgcgtg	120
Db	524	CAGAAGAGAAGGAATGCAAGAGGAGGACATTCACAGAGAGA-----	565
QY	121	tgcccttcagggaccgcttgtagcagctaaagaaggaacggagcgaggcgcaaatg	180
Db	566	-----AGGAGCCGCTTGAGCAGCTTAGAAGAGGAGCGGAGCGGAGCGCAAGATG	616
QY	181	cggagcgacagagaagagcagcgagagacagaagagcgcgagcgagcgaggaagcg	240
Db	617	CGGAGCGACAGAAAGAGACAGCGGAGCAGAAAGAGCGCGAGCGGCGGAGAGCGG	676
QY	241	cgcaagagcgagagcgcccgcaaggaaatgctgacacaaacgaataagaagagac	300
Db	677	CGCAAGAGAGCGGAGCGCGCAGGGAATGCTGACATCACCGAACATAGAGAGAGAC	736
QY	301	tacagcgacaaagttaaaagccagcacacgagatcgcaagccgcgctcgccgcgcgagag	360
Db	737	TACAGCGACAAAGTAAAGCCAGCACATGAGATCGCACCCGCGCTCGCGCGGAG	796
QY	361	cggttcgagttggagagacggccggagacgcagtaaaagaagaagaaatggaagaagagac	420
Db	797	CGGTTGAGTTGGGAGACGGCCGGAAGCCAGATAAAGAGAGAAATGAGAGAAAGGAC	856
QY	421	ctgcgtcgcagactacagacatcagcgacagcgagaggaagaccagc	468
Db	857	CTGCTGTCGCACTTACAGACATCAGCGACAGCAGAGAGAAAGACGAGC	904
RESULT	11		
LOCUS	HSU04817		
DEFINITION	HSU04817	2444 bp	mRNA
ACCESSION	U04817		PRT
VERSION	U04817.1	GI:507161	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 2444)		
JOURNAL	Xiang,J., Lantl,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.		
FEATURES	Protein kinase isoforms		
REFERENCE	J Biol Chem. 269, 15786-15794 (1994)		
SOURCE	94253170		
FEATURES	2 (bases 1 to 2444)		
REFERENCE	Kidd,V.J.		
AUTHORS	Direct Submission		
TITLE	Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's		
JOURNAL	Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St.,		
FEATURES	Memphis, TN 38101, USA		
SOURCE	Location/Qualifiers		
FEATURES	1..2444		
SOURCE	/organism="Homo sapiens"		
FEATURES	/db_xref="taxon:9606"		
SOURCE	/map="lp36-2"		
FEATURES	/cell_line="HeLa S3"		
SOURCE	/cell_type="epitheloid"		
FEATURES	/tissue_type="cervix"		
SOURCE	/clone_lib="HeLa cDNA library"		
FEATURES	/dev_stage="adult"		
SOURCE	1..2444		
FEATURES	mRNA		

Query Match	86.3%	Score 404;	DB 9;	Length 2439;
Best Local Similarity	94.2%;	Pred. No. 2.le-69;		
Matches 441; Conservative	0;	Mismatches	0;	Indels 27; Gaps 1.

gene	CDS	BASE COUNT
1. .2323	/map="1p36.3"	673 a
/gene="CDC2L2"		550 c
1104. .2287		754 g
/gene="CDC2L2"		346 t
/note="corresponding genomic sequence deposited as GenBank Accession Numbers AF080689-AF080697 and AF092426-AF092428"		
/codon_start=1		
/product="PITSLRE protein kinase beta SV8 isoform"		
/protein_id="AAC72083.1"		
/db_xref="GI:3850316"		
/translation="MKNEKKTTTSMLFQSHGSTIEIPGRVKKQRKKMGCSRVEEFOCLNRIELEGTVGVYRAKDKRTDIEIVALAKTRIKMEKEGPEITSLREINTLILKQHPNIVRREIVYGSNMDKRIIVNMYVEHDLSLMEIKRQEPLEPVSITLMQLGRAPGKHLDMN1LHRDLKTSNLLSHAGILKVGDFGLAREYSPLEKAYPPVVTQWRAPEDLTGAKESYTAIVMDVSGCIEFGELIGLQKPLFPGENSEIIOINKYFKELGTPSEKIMPGVSELPAVKKMTSFHPYVNNLRKRGALISQGGFLMKKPLTFYFGRGISMDGILKHYPFETPLPIDP5MPTFPAKSEQQGRKRGCTSPRPEEGSLGVSQGLDDDLKETGHLITTTNCGASAGAGFSLKLF"		

Query Match	86.3%	Score 404	DB 9	Length 2323
Best Local Similarity	94.2%	Pred. No. 2	le-69	
Matches 441	Conservative 0	Mismatches 0	Indels 27	Gaps 1
Qy 1	cacgaacgtctcgaaacacgacatcgagaaagaaacagataaaagctcgcgggaattggaaga	60		
Db 464	CACGAACGTCGGAAACGACATCGAAGAAACAGATTAAGCTCCGCCGGAATGGAAAGA	523		
Qy 61	cagaagagaaggaatbtygcaagggagcatcccaagagaagaagggggaatgatgctg	120		
Db 524	CAGAAGAGAAAGGAAATGGCAAGGGAGCATTCACAGAGAGAA-----	565		
Qy 121	tgcccttcaggaacgcgcttgagacgtctaaagaagcggagcggagcgaagatg	180		
Db 566	-----AAGGACCGCTTGAGACAGTTAGAAAGAAAGCGGAGACGGCAAGTG	616		
Qy 181	cggagacagcagaagagcagcggagacagaagagacgcgagcggcggagagacg	240		
Db 617	CGGGCGAGCGCAAGAGGACACGGGAGCGAGAAAGAGCCGCGCGGCGAGGACGG	676		
Qy 241	cgcaagggcgaggggcccgaggaatgctgacatacccaagataagaaagac	300		
Db 677	CGCAAGGGCGGGAGGCCCGCAGGAAATGTCTGACATCACCGAACATGAGAAAGAC	736		
Qy 301	tacagcgaacaaagtbaaaagccagccacctcggagtcgcaaccccgctcggccggcgag	360		
Db 737	TACAGCGCAAAAGTGAAGCCAGCCACACTGAGATCGCAACCCGCTCGGCGCGGGAG	796		
Qy 361	cggctgaggttgggaagcgcgcggaagccagtaaaagaagaaatgaaagaaggac	420		
Db 797	CGGTCGAGTTGGCGGAGCGCGGAGCCAGTAAAGAGAAATATGAAGAAAGGGAC	856		
Qy 421	ctgcgtcgcgaattcacagacatcagcgacagcggagaggaagacacgc	468		
Db 857	CTGCTGTCGACTTACAGGACATCAGCCACAGCGAGAGAAACCCAGC	904		
RESULT 9				
AF067521	2409 bp	mRNA	PRI	07-NOV-1998
LOCUS				
DEFINITION	Homo sapiens PIRLSLK protein kinase beta SV3 isoform (CDC12)	mRNA.		
ACCESSION	AF067521			
VERSION	AF067521.1			
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1 (bases 1 to 2409)			

AUTHORS	Gururajan, R., Lahti, J. M., Grenet, J., Easton, J., Gruber, I., Ambros, P. F. and Kidd, V. J.
TITLE	Duplication of a genomic region containing the Cdc2L1-2 and Mmp21-22 genes on human chromosome 1p36.3 and their linkage to D1Z2
JOURNAL	Genome Res. 8 (9), 929-939 (1998)
MEDLINE	98424414
REFERENCE	2 (bases 1 to 2409)
AUTHORS	Gururajan, R., Lahti, J. M., Grenet, J., Easton, J., Gruber, I., Ambros, P. and Kidd, V. J.
TITLE	Direct Submission
JOURNAL	Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
FEATURES	Location/Qualifiers
SOURCE	1..2409
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="1"
	/map="1p36.3"
gene	1..2409
	/gene="CDC2L2"
CDS	80..2383

Accession Numbers	AF080689-AF080697	and	AF092426-AF092428
/codon_start=1			
/product="PITSURE protein kinase beta SV3 isoform"			
/protein_id="NAC72086.1"			
/db_xref="GI:3850322"			
/translation="MGDEKSMWKVTIDLELQEKRRKEDEEFAEIKRLKNSDDPSK RDLSEIGELRDHCHEIRIRNSPYRREDMSDEGDSLIKIPPOGSMKEKYHNHRD EKREKKHARVKEERERKRRRREDDKRRREMRKREHRSRRRDRDLE RKREKKRMHROQKEDEKREKREARERKREARREVSANHRTREDSKVAKSH WSRPPRPRERELGDGRPVKREKREERDLSDLDIDSEKRTYSARESSAEESG GESEEEEEESEEGSTSESEEESEEESEEGSNEASEDSAESESESEES RENNHLLVVPSEFRDSESESEAEVEEGPOSALLTEGDPVPSPLIPLEIKO ELPEYLALGCRSEVFPOLIRIECTYGVYVARAKRKDEIYALKRLMEKEKGF PILSREINTILIKNOHEIVITVIREIVGSMMDKITYIMNVVENDILSLMETKOPFLP TPVYVITMIOLLRGVTKHLHDWMTIILHVDLTKLISLHLLTGDPFGIAREYSRLKAY GPVYVITOMIRAPPELLLCAKETSTRVADMSVGTCTFGLLELTPKPLPENSELIDINRYKPK ELGIPSEKIMPGSELVPKRMTESEHPYNNLKRFPALLSDGCFDLMNKEFLTPFER RISMEDLKHFFPREPLPIDPSMFPFPPAKSDEORVKNRGTISRPPEGGIGYSLDD DLKETGPHLTITNGASAGPGEFLK"			
BASE COUNT	686 a	586 c	780 g 357 t
ORIGIN			
Query Match	86.3%	Score 404;	DB 9; Length 2409;
Best Local Similarity	94.2%	Pred. No. 2,1e-69;	
Matches 441;	Conservative 0;	Mismatches 0;	Indels 27; Gaps 1;
1	cacgaactcggaaacgcacatcgagaaagaacagaataagcttcgccgggaattgggaaaga	60	
Db	434 CACGAACGTCGAAACGACATCGAAGAAACAGCATTAAGCTCGCGGAATGGGAAAGA	493	
Qy	61 cagaagagaaggaatgcaagggagcattccagggagagaagaagggaatgtagtcg	120	
Db	494 CAGAAGAAAGGAAATGCGCAAGAGAGCATTCGAGAGAGAA-----	535	
Qy	121 tgcctctcaaggacgcgtctgagacagttagaagaagacggggagcgggagcgcgaatg	180	
Db	536 -----AGGGAGCCGCTTGAGCAGCTTAAAGAAAGAACCGGAGCGGAGCCGCAAGATG	586	
Qy	181 cgggagcagcagaagagacacgcgggaagaagaagacgcgagcggcgggagagaagcgg	240	
Db	587 CGGGAGCGAGCAAGAGACAGGGGAGGAGAAAGGAGCCCGCGAGCGGAGAGAGCGG	646	
Qy	241 cgcgaagagcgggagcccgccaggggaatgtctgacataccggaacgaatgagaagagc	300	
Db	647 CGCAAGGAGCGGAGAGCCCGCAGGGAAGTGTCTGCATCAACCGAAGATGAGAGAGAC	706	
Qy	301 tacagcgacaaatgtgaaagccaagccatcggagtcgcaagccgcgcttcgycgcgcgggag	360	
Db	707 TACACCGACAAAGTGAAGGACGACCACTGAGGATGCGACCCGCTTCGGCCGCGGGAG	766	

	Db	431	CGGGAGCAGCAACAAAGAGCAGCGGGAGCAGAAGACC	GCGACGGCGGGCGGAGCAGCGG	490
	OY	241	cgcgaagcagcgaggccgcccacagggaaagtgtctcacatcaaccgaacgatgagaagggac		300
	Db	491	CGCAAGAGAGCGGAGGCCCGCAGGGAAGTGTCTGCATCATCACGAACGATGAGAGAGNC		550
	OY	301	tacagcagacaagtataaagccagccacctgtagtgcacgcccgcttcgagccgcgcgaggaag		360
	Db	551	TACACCGCACAAGGTATAAAGCCAGCATCTGACTTCACACCCCCTTCGCGCGCGGGAGG		610
	OY	361	cggttccgagtttgaggagacgcgcggaagccagtaaacagaagatgaaatgaaagaggac		420
	Db	611	CGGTTCCGAGTTTGAGAGACGCGCCGGAACCACTTAAGAAGACAAAATGCAAGAAAGGAC		670
	OY	421	ctgcgtccgacttacagacatcagcagacagcagaggaagaccagc		468
	Db	671	CTGCTGTCGCACTTACAGGACATCAAGCAGCAGAGGAAGAACACAGC		718
RESULT	7				
LOCUS	AF067520	2436 bp	mRNA	PRI	07-NOV-1998
DEFINITION	Homo sapiens PITSRE protein kinase beta SV2 isoform (CDC2L2) mRNA,				
ACCESSION	AF067520				
VERSION	AF067520.1	GI:3850319			
KEYWORDS	. human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;				
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 2436)				
TITLE	Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P.F. and Kidd,V.J.				
JOURNAL MEDLINE REFERENCE AUTHORS	Duplication of a genomic region containing the Cdc2L1-2 and MHP21-22 genes on human chromosome 1p36.3 and their linkage to D1Z2 Genome Res. 8 (9), 929-939 (1998)				
	98424414				
	2 (bases 1 to 2436)				
	Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P. and Kidd,V.J.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA				
FEATURES	Location/Qualifiers				
source	1..2436				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/chromosome="1"				
	/map="1p36.3"				
	1..2436				
gene	/gene="CDC2L2"				
	80..2410				
CDS	/note="Corresponding genomic sequence deposited as GenBank accession numbers AF080689-AF080697 and AF092426-AF092428"				
	/codon_start=1				
	/product="PITSRE protein kinase beta SV2 isoform"				
	/protein_id="AAC72085.1"				
	/db_xref="GI:3850320"				
	/translation="MGDEKDSWKYTLDELILQEKRRRKEQEKAELRLNKSDDRSKR RSLSEGLRDHCMEITIRNSPYRREDSDMEDRGEDDSLAIKPPQMSWEKVHHRDD EKREKKHARVAREHERRRKHREDODKARREROKRREMARESRREANGDFCL FRODLQLEREREKRRMRKQROKROROKREERARERRKREARRRVSAHHRTRMD YSKVKASHMSRSRP RP RP RP RP FELGDGRKPKYEKMEERDLSLDIISDEKRTSSA ESSSAESGSGESEEEEEESEEGSTSESSESESEEEEEETSGNSEASQSAEVS EEMSEDEBERENENHLVPESEDRDGSSESEEBEYGVTGYVRADKDTDEIVALKRL ALPIETLKQELPKYLPAIGCRSVEEFOCLNRIIEGTGYGVYRAKDKTDEIVALKRL KMEKEKGFPITSLREINTLIKQHNPNIIVREIVGSSNDKIYIVANNVEHDIKSLM ETMKOPLPREVKTIIMQLARGYHLDNNMLRHDLTKSNLLSHAGITAKVGDPGLAR EYGSPLAIVPVAVVTWYRAVELLGKEYSTADVMSVCIIFELLTQKPLPFGNSE IDIDNKFKELGTPSEKITWPGYSILPVVKKTWTSEEHYNLIRKFFGLISDQGDLANN				

KFILYFGRRRISADGEKHEFRFEMPLPIDPSMPTPMRPAKSDQRYVKRIGTSPRPPEGG									
LGIYQLDCHDDDKETGFEHLITTTNOCGASAPGFSIFK"									
BASE COUNT	690 a	591 c	789 g	366 t					
ORIGIN									
Query Match	99.3%	Score 464.8:	DB 9:	Length 2436:					
Best Local Similarity	99.6%:	Pred. No. 3e-81:							
Matches 466:	Conservative	0:	Mismatches	2:	Indels	0:	Gaps	0:	
OY	1	cagcagacgtcggaaacacgacatcgagaagaacagagataaagctcgccgggaatgggaaaga	60						
DB	434	CACGAAACGTGGAAACGACATCGAAGACACGGATTAAGCTCGCCGGGAATGGAAAGA	493						
OY	61	cagaagagaaggaatggcgaagggagcatctccagagagagaagaaggggaaatgatgctg	120						
DB	494	CAGAAGAGAGGGAATGGCAGGGAGCATCTCCAGAGAGAGAAAGGGGATGATGGCTTT	553						
OY	121	tgctcttcgaaggaccgcttggagcaagttagaagaaggaacgggagccggagcgcgaatg	180						
DB	554	TGCTCTTTCGAGGACCGCTTTGGAGCAGTTAGAAAGAAAGCCGGACCGGAGCGCAAGATG	613						
OY	181	cgaggagcgaacgaagaagagcggagacgaagaaggagcgcgagccggcgaggagaacg	240						
DB	614	CGGAGCACCAAGAGAGCAGCGGAGACAGAAAGAGCGGACCGCGGGGAGAGAGCGG	673						
OY	241	cgcagaagcggaggagcccgccaggggaagtgtctgtacatcacgaacgatgtagagaggac	300						
DB	674	CCCAAGAGACGGGAGGCGCCGACGAGAGTGTCTGCACATCACGAAACGATGAGAGAGAC	733						
OY	301	tacagcgcgaagaagtgaagccagccactgtgagtcgcagccgcgctcgccgcgcggag	360						
DB	734	TACACGCAAGAAAGTAAAGCCAGCCACTGTGATCTCCAGCCGCCCTCGCGCGGGGAG	793						
OY	361	cggttcgagttggaggacggcgaggagccgaftaaagaagaagaataatggaaagaaggac	420						
DB	794	CGGTCGAGTTGGGAGACGCGCCGAGAGCCAGTAAAGAGAGAAATGAGAAAGAGGAC	853						
OY	421	ctgcgtccgcactcacagacatcagcgacagcgaacgagaggaagaccagc	468						
DB	854	CTGCTGCTCGGACTTACAGGACATCAGCGACACGAGAGAGAACACCAAC	901						
RESULT	8								
LOCUS	AF067518	2323 bp	mRNA	PRI	07-NOV-1998				
DEFINITION	Homo sapiens PIMSLRE protein kinase beta SV8 isoform (CC212) mRNA.								
ACCESSION	complete cds.								
VERSION	AF067518								
KEYWORDS	AF067518.1 GI:3850315								
SOURCE	human.								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.								
	1 (bases 1 to 2323)								
	Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,								
	Ambros,P.F. and Kidd,J.V.								
TITLE	Duplication of a genomic region containing the Cdc2L1-2 and								
	MM21-22 genes on human chromosome 1p36.3 and their linkage to D12Z								
JOURNAL	Genome Res. 8 (9), 929-939 (1998)								
MEDLINE	98424414								
REFERENCE	2 (bases 1 to 2323)								
AUTHORS	Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,								
	Ambros,P. and Kidd,J.V.								
TITLE	Direct Submission								
JOURNAL	Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's								
	Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA								
FEATURES	Location/Qualifiers								
SOURCE	1..2323								
	/organism="Homo sapiens"								
	/db_xref="taxon:9606"								
	/chromosome="1"								

JOURNAL Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St., Memphis, TN 38101, USA

FEATURES
source

1. .2471
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="1p36-2"
/cell_line="Hela S3"
/cell_type="epitheloid"
/tissue_type="cervix"
/clone_lib="Hela CDNA library"
/dev_stage="adult"
1. .2471
112. .2445
/codon_start=1
/product="PITSLRE alpha 2-2"
/protein_id="AA19582.1"
/db_xref="GI:307160"
/translation="MGDKDSMKYKTTIDELIQEKRRKEQEKAIKRLKNSDDRSK
RDSLEGEIRDHCMETIRNSPYRREDSMERGEEDSLAIKPPQMSMKKRVHHRKD
EKRRKKHARVKKERHEHRRKRHREEDOKRAREEROKRREMARHESRRGNDYVC
LFDRLEOLEKREBERKREDOOKROREOKRERAREERKEREAREVSAHHRMRE
DYSOKVKAHMSRPPRRPRPRPRPRPRPRPRPRPRPRPRPRPRPRPRPRPRPRPR
AESSAESGSGSEEEEEESEEESEEESEEESEEESEEESEEESEEESEEESEEESE
SEEMSEDEERENENHLLVPESEPRDROSGSESEEEVGGTPOSSALTLTGIDVPDS
PALSPLEIKOELPKYLPALOGCRSVEEEOCNRIEEGYVYRAKDKTDEIVALKR
LKMEKEGEPITSLREINTILKAQHPIVIRLEIYSNMDKIYIVANVYEHDLKSL
METMKOPFLPGEVKTIIMIOILRGVYKHLHDNMILHRDLKTSNLLSHAGILVGEGLA
REYSGPLKATPVVYVITWYRAPELLIGAEKYSTAVDMSSVCIFGELLTOKPLRPGKS
ELDOIINKYFKDILGTPSEKIWPQYSELIPAVKKMTTSRHPYNNLRKRFGLSLDQFDLM
NKEIYFPGRRISAEQGLKHEFRETPLPLDPSMPEYTPAPSEDOORVAKRGTSPPEPS
GIGYSQLDDDLKERTGHEFHLTTTNOGASAGPFSILKF"

BASE COUNT 697 a 601 c 809 g 364 t
ORIGIN

Query Match 100.0%; Score 468; DB 9; Length 2471;
Best Local Similarity 100.0%; Pred. No. 7e-82;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cagaagcgtcggaagacatcgagaagacagatgaagtcgcgcgggaatggaaga 60
DB 469 CACGAACGTCGGAAACGACATCGAGAGAACAGATTAAGCTGCGGGAATGGGAAAGA 528
QY 61 cagaagagaaggaatggaaggaagcattcgaagagaagaaggggaatgagtcgtg 120
DB 529 CAGAAGAGAAAGGAATGGCAAGGAGCATTCACAGAGAGAAAGGGGGAATGTCGCGTG 588
QY 121 tgcctcttcaggagccgcttgagcaggttagaagaagacgaggagcggaagatg 180
DB 589 TGCCTCTTTCAGGAGCCGCTTGAGAGATTAGAAAGAACCGGAGCGGAGCAAGATG 648
QY 181 cggagacagcagaagagcagcgaggagcagaagagcgagcgcgaggagcgagcg 240
DB 649 CGGGAGCAGCAGAGAGAGAGCAGCGGAGCAAGAGAGCGGAGCGGCGGAGAGCGG 708
QY 241 cgaagagcgaggagccgcgacgaagagatgctcgcacatcaccgaacgaatgagaagagac 300
DB 709 CGCAAGGAGCGGAGCGCGCAGGAGGAGTGTCTGCACATCACCGACGATGAGAGAGGAC 768
QY 301 tacaagcagaagaatgaaagccacatcgagtcgagcccgcttcgcccgcgcgagag 360
DB 769 TACAGCCACAAAGTGAAGCCACGCTGAGTCGACGCCCGCTTGCGCGCGCGGAG 828
QY 361 cggcttcagtcgagagcgcggaagccagtaaaagaagaagaatggaagaagaagagac 420
DB 829 CGGTTCAGTTGGGAGAGCGCGGAGGCCAGTAAAGAAAGAAAGAAATGGAAGAAAGGAC 888
QY 421 ctgcgtctcgaactacagaacatcagcgacgacgagaggaagaacagcagc 468
DB 889 CTGCTGTCCGACTTACAGGACATCAGCAGCGAGAGAGAGAGACGACG 936

RESULT 6
AF067525 2329 bp mRNA PRI 07-NOV-1998
LOCUS
DEFINITION Homo sapiens PITSLRE protein kinase beta SV13 isoform (CDC2L2)
mRNA, complete cds.

ACCESSION AF067525.1 GI:3850329
VERSION
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Gururajan,R., Lahit,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P.F. and Kidd,V.J.
TITLE Duplication of a genomic region containing the Cdc2L1-2 and
MMP21-22 genes on human chromosome 1p36.3 and their linkage to D122
Genome Res. 8 (9), 929-939 (1998)

JOURNAL MEDLINE
REFERENCE
AUTHORS Gururajan,R., Lahit,J.M., Grenet,J., Easton,J., Gruber,I.,
Ambros,P. and Kidd,V.J.
TITLE Direct Submission
JOURNAL Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's
Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA

FEATURES
source
1. .2329
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
1. .2329
/gene="CDC2L2"
68. .1654
/gene="CDC2L2"
/note="corresponding genomic sequence deposited as Genbank
Accession Numbers AF080689-AF080697 and AF092426-AF092428"

BASE COUNT 634 a 605 c 738 g 352 t
ORIGIN

Query Match 99.3%; Score 464.8; DB 9; Length 2329;
Best Local Similarity 99.6%; Pred. No. 3e-81;
Matches 466; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 cagaagcgtcggaagacatcgagaagacagatgaagtcgcgcgggaatggaaga 60
DB 251 CACGAACGTCGGAAACGACATCGAGAGAACAGATTAAGCTGCCCGGGAATGGGAAAGA 310
QY 61 cagaagagaaggaatggaaggaagcattcgaagagaagaaggggaatgagtcgtg 120
DB 311 CAGAAGAGAAAGGAATGGCAAGGAGCATTCAGAGAGAGAAAGGGGGAATGATGAGCTTT 370
QY 121 tgcctcttcaggagccgcttgagcaggttagaagaagagcgaggagcggaagatg 180
DB 371 TGCCTCTTTCAGGAGCCGCTTGAGAGATTAGAAAGAGCGGAGCGGAGCAAGATG 430
QY 181 cggagacagcagaagagcagcgaggagcagaagagcgagcgcgaggagagcgagcg 240

Query Match	100.0%	Score 468;	DB 6;	Length 660;
Best Local Similarity	100.0%;	Pred. No. 8,8e-82;		
Matches 468; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

mRNA	1..2465
gene	112..2439
CDS	/gene="p58DC2L1" 112..2439 /gene="p58DC2L1"

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:54:31 ; Search time 2399.05 Seconds
(without alignments)
3218.224 Million cell updates/sec

Title: US-09-915-060-5

Perfect score: 468
Sequence: 1 caccgaactcggaaacgaca.....acagcgaagagaccagc 468

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_om: *
20: em_or: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vi: *
30: em_htgo_hum: *
31: em_htgo_in: *
32: em_htgo_rod: *
33: em_htgo_hum: *
34: em_htg_in: *
35: em_htg_rod: *
36: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	468	100.0	468	6	AX033424
2	468	100.0	660	6	AX033425
3	468	100.0	2465	6	HSU07705
4	468	100.0	2471	6	AX033422
5	468	100.0	2471	6	HSU04816
6	464.8	99.3	2329	9	AF067525
7	464.8	99.3	2436	9	AF067520
8	404	86.3	2323	9	AF067518
9	404	86.3	2409	9	AF067521
10	404	86.3	2439	9	AF067519
11	404	86.3	2444	9	HSU04817
12	404	86.3	2471	9	HSU07704
13	404	86.3	2477	9	HSU04824
14	404	86.3	2486	9	AF067512
15	404	86.3	2500	9	AF067514
16	404	86.3	2533	9	AF067516
17	355	75.9	2362	9	AF067523
18	355	75.9	2448	9	AF067522
19	355	75.9	2525	9	AF067515
20	348.6	74.5	2544	9	AK000081
21	296.8	63.4	3161	10	MUSCDEPK
22	203.4	43.5	2340	9	HSU04818
23	203.4	43.5	2349	9	AF067517
24	181	38.7	1715	9	AF174497
25	166	35.5	110608	9	HS28383
26	165	34.6	297	9	CDC2LIS06
27	161.8	33.3	297	9	CDC2LIS06
28	151	32.3	1936	10	MUSP58GTA
29	138.4	29.6	712	9	AF067524
30	127.4	27.2	532	9	CDC2LIS07
31	127.4	27.2	533	9	CDC2LIS07
32	105.4	22.5	539	9	CDC2LIS05
33	105.4	22.5	539	9	CDC2LIS05
34	105	22.4	2226	9	AF067513
35	100	21.4	551	9	AF067527
36	100	21.4	700	9	AF067529
37	99.4	21.2	7218	6	AF067526
38	98	20.9	579	9	AF067528
39	98	20.9	728	9	AF067528
40	79	16.9	480	9	CDC2LIS08
41	79	16.9	592	9	CDC2LIS08
42	72.4	15.5	484	5	CPTSLRE02
43	67.2	14.4	178022	2	AC018348
44	67.2	14.4	179880	2	AC091304
45	66.2	14.1	238688	2	AC079818

ALIGNMENTS

RESULT 1
AX033424
LOCUS AX033424 468 bp DNA
DEFINITION Sequence 5 from Patent WO0044896.
ACCESSION AX033424
VERSION AX033424.1 GI:10280185
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 468)
AUTHORS Beyaert, R. and Cornelis, S.
TITLE Internal ribosome entry site (IRES), vector containing same and
uses thereof
JOURNAL Patent: WO 0044896-A 5 03-AUG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS
SIGRID (BE)
FEATURES
source location/Qualifiers
1..468


```

Db      334 GCAGATATCAGGCTCAGGTTCTGGAAGAGGAGGACGAGGAACGAGAAACGAGGAAGACGA 393
Oy      538 gaggaaggaaggaagcaccagtcgaagaatcagaggaaggaaggaaga-----584
Db      394 ACGAAGAACGGGAGACACACAGTGAACGAACTCAGAGGAAGAAACGAGGAAGGAAGAGA 453
Oy      585 -----ggaaggaaggaagacgcgcagcaactctggaaggaagcaacgaagcgtcgcgaag 640
Db      454 GGACGGAAGAGAGGAGAGACTGGAGACCACTCTGAGAGAGCGCTCTGCACAGTCCGACAGAAG 513
Oy      641 aagtaagtgaggaagaatg 660
Db      514 AAGTCAGTCATGATGAGGAATG 533

RESULT 14
Bg943585 LOCUS      326 bp      mRNA      EST      11-JUN-2001
DEFINITION ax39f04.x1 Proliferating Human Erythroid Cells (LCB:ax library)
ACCESSION  Bg943585
VERSION    Bg943585
KEYWORDS   Homo sapiens cDNA clone ax39f04 random, mRNA sequence.
SOURCE     Ery.
ORGANISM   human.
REFERENCE  1.
AUTHORS   Gulin,A.N., Mjoroge,J.M., Bouffard,G.G. and Miller,J.L.
TITLE     Gene expression in proliferating human erythroid cells
JOURNAL   Genomics 59 (2), 168-177 (1999)
MEDLINE   99339981
COMMENT    Contact: Jeffery L. Miller
          Laboratory of Chemical Biology
          National Institute of Diabetes and Digestive and Kidney Diseases
          Building 10, Room 9B17, National Institutes of Health, Bethesda, MD
          20892, USA
          Tel.: 301 402 2373
          Fax: 301 435 5148
          Email: jlm7fe@nih.gov
          DNA Sequencing and analyses by National Institutes of Health
          Intramural Sequencing Center (NISC).
          Plate: 39. row: f column: 04
          Seq primer: 21M13 forward primer (ABI).
          Location/Qualifiers
            1..326
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="ax39f04"
               /clone_lib="Proliferating Human Erythroid Cells (LCB:ax
               library)"
               /sex="unknown"
               /tissue_type="blood"
               /cell_type="Erythroid Cells"
               /cell_line="Primary Culture of Peripheral Blood
               Mononuclear Cells"
               /dev_stage="progenitor; EPO responsive CD71++++"
               /lab_host="SOLR"
               /note="Organ: Blood, Vector: Lambda ZAP II; Site1: EcoRI;
               Site2: EcoRI; 65,000 proliferating erythroid cells from
               the buffy coat of a blood donation were obtained by flow
               cytometric separation after a 5-day culture period in the
               presence of erythropoietin. Total RNA was purified from
               the sorted cell population using TRIzol reagent. RNA (0.3
               ug) was converted into double stranded cDNA using
               Clontech's Capfinder cDNA Library Construction Kit
               (Clontech) according to the manufacturer's protocol and
               cloned into EcoRI digested lambda Zap II vector
               (Stratagene). The phage library was amplified once prior
               to in vivo excision in SOLR cells. Individual colonies
               were grown, and the cDNA inserts were sequenced in high
               throughput (NIH Intramural sequencing center
               http://www.nisc.nih.gov/)."

```

BASE COUNT	109 a	67 c	128 g	22 t
ORIGIN				
Query Match	41.2% ; Score 271.6 ; DB 11; Length 326;			
Best Local Similarity	96.9% ; Pred. No. 2,36-35;			
Matches 277; Conservative	0; Mismatches 9; Indels 0; Gaps 0;			
Qy	130 agggaccgcgtcttgagcagtttaagaaggaagcgggagcgcggaagatgcggagacag 189			
Db	35 AGGAGCCGCTTTGAGAGCATTTAGAAAGAAAGCGGAGCCGGAGCCAGATGCGGAGACAG 94			
Qy	190 caagaagaagcaagcgagggagacagaagagcgcgagcgcgagcgagagagcgcgcaagagag 249			
Db	95 CAGAGGAGGACGAGCGGGAGCCAGAAAGGAGCGCGAGCGCGGAGGAGAGAGCGCGGAGAG 154			
Qy	250 cggagagcgccgcgagggaggaagtgtctgcacatcacgcgaagatgagaagagagactcaagagac 309			
Db	155 CGGAGAGCCCGCAGAGGAAGTGTCTGCACATCTCACCAAGAGATGAGAAGAGGACTACAGCGAC 214			
Qy	310 aagaatgaagccacgcacacgtcgaggtcgcaacgcgcgcgttcgcccgcgcgagagcggttcgaag 369			
Db	215 AAATGTAAACCCAGCCACACTGTGAGACTCGCAGCCGCGCTTCGGCCGCGCGGAGAGCGGTTTGAG 274			
Qy	370 ttgggagacgagccggaagccagttaaagaagaagaataatgtgaagaa 415			
Db	275 TTGGGAGACGCGCGGAGAACGACAGAAAAAATAAAAAAATAAAAAA 320			
RESULT 15				
BG969359	BG969359 758 bp mRNA EST 12-JUN-2001			
LOCUS	60283684.1F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4971155			
DEFINITION	5', mRNA sequence.			
VERSION	BG969359			
KEYWORDS	BG969359.1 GI:14356996			
SOURCE	EST.			
ORGANISM	house mouse.			
REFERENCE	Mus musculus			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
TITLE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
JOURNAL	1 (Phases 1 to 758)			
COMMENT	NIH-MGC http://mgc.nci.nih.gov/.			
	National Institutes of Health, Mammalian Gene Collection (MGC)			
	Unpublished (1999)			
	Contact: Robert Strausberg, Ph.D.			
	Email: cgabs-r@mail.nih.gov			
	Tissue Procurement: Jeffrey E. Green, M.D.			
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)			
	cDNA Library Preparation: Life Technologies, Inc.			
	DNA Sequencing by: Incyte Genomics, Inc.			
	Clone distribution: MGC clone distribution Information can be			
	found through the I.M.A.G.E. Consortium/LLNL at:			
	http://image.llnl.gov			
	Plate: LLAM0956 row: a column: 12			
	High quality sequence start: 3			
	High quality sequence stop: 731.			
FEATURES	Location/Qualifiers			
source	1..758			
	/organism="Mus musculus"			
	/strain="FVB/N"			
	/db_xref="taxon:10090"			
	/clone="IMAGE:4971155"			
	/clone_1lb="NCI_CGAP_Kid14"			
	/lab_host="DH10B (TI phage-resistant)"			
	/note="Organ: Kidney; Vector: pCMV-SPORT6; Site_1: NotI;			
	Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.			
	Average insert size 1.75 kb. Constructed by Life			
	Technologies. Note: this is a NCI_CGAP Library. "			
BASE COUNT	254 a	144 c	254 g	106 t
ORIGIN				

REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Rodentia; Sclauognathi; Muridae; Murinae; Mus.
TITLE	1 (bases 1 to 830)
JOURNAL	yahyawi,M., Hennig,S., Neihardt,L., Radelof,U., Hermann,B.G., Lehnach,H. and O'Brien,J. Detection of a high number of novel genes in a 9-day mouse embryo cDNA library normalised by oligonucleotide fingerprinting Unpublished (2001)
COMMENT	Contact: Hennig S laboratory 123, dept.Lehnach Max-Planck-Institut fuer Molekulare Genetik Inhestr.63-73, D-14195 Berlin, Germany Tel.: +49 30 8413 1612 Fax: +49 30 8413 1380 Email: hennigemolgen.mpg.de EST's are made from clones being representatives of clone clusters Clone clusters were calculated from oligonucleotide fingerprints. PCR primers FORWARD: 5'-GAGCTATTCCAGAAGTAGTGA-3' BACKWARD: 5'-TAATACGACTCACTAATAGG-3' Seq primer: 5'-ATTTAGGTGACACTTAG-3' High quality sequence stop: 830.
FEATURES	Location/Qualifiers 1..830
source	/organism="Mus musculus" /db_xref="taxon:10090" /clone="ICRFp522I0930" /clone_lib="Mouse 9-day fetus cDNA library ICRFp522" /tissue_type="whole embryo" /dev_stage="embryonic 9-day" /lab_host="E.coli, XLI blue" /note="vector: pSVSPORT1; site_1: NotI; site_2: SalI; library preparation by oligo_dT priming of RNA. Clones can be ordered from the Resource Center in Berlin, http://www.rzpd.de "
BASE COUNT	268 a 161 c 271 g 125 t 5 others
ORIGIN	
Query Match	43.4%; Score 286.4; DB 11; Length 830;
Best Local Similarity	80.1%; Pred. No.8.4e-38;
Matches	367; Conservative 0; Mismatches 76; Indels 15; Gaps 2;
Db	218 gccagcgccgagcgaggacgcggcaagaagccgccgaagatgtctgcac 277
Db	44 GGGAACGGAGGGCCAAGAGGCCCGCCAAAGAGAGAACCCGCTAGGGGAAGCTCTGCAC 103
OY	278 ataccggaacgcttgtagagagacttcacgcgaagaagttaaagccaagtcagtgcga 337
Db	104 ATCACCTTACCCTTGAGGGAGAGTCAGTGTGATTAAGGGAGAGTTGGCCACTGGACCGGA 163
OY	338 gccgcctcgcgcgcgcgcgcgcgcggttcgagttlvggagacgcgcgcgaagccagtaaaag 397
Db	164 GCCCCTTGAGGCCACCAAGAGAGCGCTTTGAGATGGAGACAACCGGAAGCCAGTAAAG 223
OY	398 aagagaatatgtgaagaagagaccctgcctgcaccttcaggaactcaagcagcagaga 457
Db	224 AAGAGAAGGTGGAAGAGAGACCTTGTTGTCAAGACCTTCAGACACTTAAGTGCACACCGAGA 283
OY	458 ggaagaacagctcgcgcgcgactcctcgtaagcagaattagagcttcagttcagaagaag 517
Db	284 GSAAAACCAAGCTCAAGCTTCTTCATATACGACGATTCAGGCTCAGGTTCTGAAGGGAGG 343
OY	518 aggaaggaaggaagaag--gagagaagaagaagagaccagtygaaatcaagaagaag 574
Db	344 AGGAGGAAGAAGAGGAAGAAGAAAGAAAGGAGACCAACAGTGAAGATCATATGGAAG 403
OY	575 aagagaaggaaggaagag-----aggagaaccgcgcgcgaactcttaagtagagcat 622
Db	404 AAGAGGAAGAAGAGGAAGAAGAAAGAGGAGACTGGGAGCAACTTGAGGAGGCGCT 463
OY	623 cagaagcagctgcgcgaagaagtaagttgaggaagaatg 660

[illegible]

```
/clone_lib="NIH_MGC_87"
/tissue_type="mammary adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: breast; Vector: PCMV-SpOre6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.383 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT      272 a      165 c      270 g      97 t
ORIGIN

Query Match      47.6%; Score 314.2; DB 11; Length 804;
Best Local Similarity 84.5%; Pred. No. 2.4e-42;
Matches 388; Conservative 0; Mismatches 43; Indels 28; Gaps 2;

QY 1 cagcaacgtcggaaacacatcgaagaacacagatataagcccgccgggaatgggaaga 60
    |||||||
Db 358 CACGAACGTGGAAACGGCATCGAGAAACAGATAAAGCTCGCCGGGAATGGGAAGA 417
    |||||||

QY 61 cagaagagaagaaatggcaagagacatcccaagagaagaaggggaatgatgcytg 120
    |||||||
Db 418 CACAAGAGAAGGAGATGGCAAGGAGCATTCACAGAGAGAA----- 459
    |||||||

QY 121 tgcctctcagggaccgcttgagacgttaagaagaacggagagcggaagcgcaatg 180
    |||||||
Db 460 -----AGGGACCGCTTGGAGCAGTTAGAAAGGAACGGGAGCGCAGCGCAAGATG 510
    |||||||

QY 181 cggagacagcagaagagcggagcggagcgaagaagcggagcggcggaagagcg 240
    |||||||
Db 511 CGGAGACAGCAGAGAGAGCGCGGAGCAGAAAGAGAGCGGAGGGGCGGAGAGCGG 570
    |||||||

QY 241 cgaagagcggagagcccgcaaggaagtgtctgcacatcacgaacgaatgagaagag 300
    |||||||
Db 571 CGCAAGAGAGCGGAGCGCCGACAGCAAGTGTCTGCACATCACCGAAGCATGAGAGAGAG 630
    |||||||

QY 301 tacagcgaacaagtgaagcccaagcaactgtgagtcgacgccgcctcgccgcggagag 360
    |||||||
Db 631 TACAGGACAAATGAAAGCACCGACCTGAGTCGACGCCCGCTCGCGCGCGGAG 690
    |||||||

QY 361 cggcttgagctggagagcgg-ccggaagccagtaaaagaagagaataatgaagaagga 419
    |||||||
Db 691 CGGTTGAGTTGGGAGACGGTCCGGAAAGCCAGTAAAAAAGCAGCACACTGCAGCAAG 750
    |||||||

QY 420 cctgctgtccgaactacagacatcagcgacagcgagag 458
    |||||
Db 751 GGACGTGCCGCGCTCGACTTTCACAGGAGCATACGACGAG 789
    |||||

RESULT 11
BI018825/c 493 bp mRNA EST 14-JUN-2001
LOCUS BI018825 IL3-MT0267-050101-435-H02 MT0267 Homo sapiens CDNA, mRNA sequence.
DEFINITION BI018825
ACCESSION BI018825
VERSION BI018825.1 GI:14425455
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 493)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL MEDLINE
COMMENT 20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
```

```
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=IL3&t2=IL3-MT0267-
050101-435-H02&t3=2001-01-05&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 13
High quality sequence stop: 283.
Location/Qualifiers
1. 493
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MT0267"
/dev_stage="Adult"
/notes="Organ: marrow; Vector: puc18; Site: 1: SmaI; Site: 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

BASE COUNT      66 a      178 c      90 g      159 t
ORIGIN

Query Match      47.1%; Score 310.6; DB 11; Length 493;
Best Local Similarity 91.2%; Pred. No. 9.5e-42;
Matches 344; Conservative 0; Mismatches 24; Indels 9; Gaps 1;

QY 293 gagaagactacagcagcaaatgaaagccacagccactgagtcgagccgcctgcgcgc 352
    |||||||
Db 493 GAGAGATTACAGCAGCAAAAGTGAAGGACGATGGAGTCGCCAGCCGCTTCGAGAGC 434
    |||||||

QY 353 cgcggagacggttcgagttggagagcggccggaagccagtaaaagaagaatggaag 412
    |||||||
Db 433 CGCGGAGACGGTTCGAGTTGGGAGAGCGCCGCAAGCCAGTAAAGAGAGCAAAATGGAAG 374
    |||||||

QY 413 aaagggaccctgtctgcacttacagagacatcagcgagagagagaaagaaacccgcgcg 472
    |||||||
Db 373 AAAGGACCTCTGTCCTTCACAGGCGCATTTGACAGGAGAGAGAGAGCCAGCTCGG 314
    |||||||

QY 473 ccgaatcctctcagcagaatcaggtctcttgagaaagagagagagagagag 532
    |||||||
Db 313 CCGAGTCTCTGCTCATTTTATCAGGTCCAGGTTCTGAGCAATTAACAGAGAGAGAGAG 254
    |||||||

QY 533 aggaagagagagagagagacccagtgagaatca-----gagagagagagagag 583
    |||||||
Db 253 AGGAGAGAGAGAGAGAGGAGCACCAGTGAAGATCAGAGAGAGAGAGAGAGAGAGAGAG 194
    |||||||

QY 584 aggaagagagagagagagccgcaagcaactctgagagagagcatcagagcagcttcgcgaaag 643
    |||||||
Db 193 AGGAGAGAGAGAGAGAGCCGCGCAACTCTGAGAGGAGCATGAGACAGTCTCGAGAAAG 134
    |||||||

QY 644 taagtgaagaagaatg 660
    |||||||
Db 133 TAAGTGAGGAGAGAAATG 117
    |||||||

RESULT 12
BF607177 830 bp mRNA EST 01-APR-2001
LOCUS BF607177
DEFINITION MYL_000119 Mouse 9-day fetus CDNA library ICRP522 Mus musculus
CDNA clone ICRFP52210930 5', mRNA sequence.
ACCESSION BF607177
VERSION BF607177.1 GI:13503669
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
```

Best Local Similarity 91.3%; Pred. No. 5.3e-46;
Matches 379; Conservative 0; Mismatches 9; Indels 27; Gaps 1;

```
QY 1 caggaacgtcggaagacacatcggaagaagataaagtcgcgcggaatggaaga 60
Db 508 CACGAACGTCGGAACGGCATCGAAGAACAGATTAAGCTCGCGGAATGGGAAGA 567
QY 61 cagaagaagaagaatggaagagacatccagagagaagaagggaatgatgagcgtg 120
Db 568 CAGAAGAGAAAGGAGATGGCAAGGAGCATTCACGAGAGAGA----- 609
QY 121 tgccttcaaggaccgcttggaagcagttagaagaagcggagacggagacgaatg 180
Db 610 -----ACGGACCCGCTTGACCACTTAGAAAGCAAGCGGAGCGGAGCCAGATG 660
QY 181 cggagagacagcagaagcagcggagacagaagaagcggagcggcgaggagagacg 240
Db 661 CGGAAGCGACGAAGAGACCGCGGACAGAAAGGAGCGCGGCGGCGAGAGACGG 720
QY 241 cgaagaagcggagagccgcgaggaagtgctcacatcaacgaacgatgagagagac 300
Db 721 CGCAAGAGCGCGGAGGCCCGCAGGAGTGTCTCATCATCACCGAAGATGAGAGAG 780
QY 301 tacagcgacaaagtgaagccagccactggaagtcgacgcccgcctcgccgcggagag 360
Db 781 TACAGCGACAAAGTGAAGCCAGCCACTGAGTGCAGCCCGCTCGCGCCGCGGAG 840
QY 361 cggctcgaattggagagcggagacagtaagaagaagaatggaagaagaa 415
Db 841 CGGTTCCAGTTGGGAGACGGCGGAGCCAGTAAATAAAAAAAAAAAAAA 895

RESULT 9
AI580044/c 592 bp mRNA EST 14-DEC-1999
LOCUS tg45908.x1 NCI-CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2211806 3,
DEFINITION similar to TR:061399 Q61399 CELL DIVISION CYCLE 2-LIKE 2; contains
ACCESSION AI580044
VERSION AI580044.1 GI:4564420
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 592)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmett-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbfp/image/image.html
Insert Length: 1082 Std Error: 0.00
Seq primer: -40UP from G1bco
High quality sequence stop: 405
POLYA=No.
```

FEATURES

source location/Qualifiers

1..592

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2211806"

/clone_lib="NCI CGAP Ut1"

/tissue_type="well-differentiated endometrial adenocarcinoma, 7 pooled tumors"

/lab_host="DH10B"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014"

```
BASE COUNT 58 a 212 c 115 g 207 t
ORIGIN
Query Match 50.4%; Score 332.8; DB 10; Length 592;
Best Local Similarity 92.8%; Pred. No. 2.2e-45;
Matches 371; Conservative 0; Mismatches 2; Indels 27; Gaps 1;
QY 1 caggaacgtcggaagacacatcggaagaagataaagtcgcgcggaatggaaga 60
Db 373 CACGAACGTCGGAACGGCATCGAAGAACAGATTAAGCTCGCGGAATGGGAAGA 314
QY 61 cagaagaagaagaatggaagagacatccagagagaagaagggaatgatgagcgtg 120
Db 313 CACAAGAGAAGGAGATGCAAGGAGCATTCAGAGAGAA----- 270
QY 121 tgccttcaaggaccgcttggaagcagttagaagaagcggagacggagacgaatg 180
Db 271 -----AGGACCGCTTGAGCACTTAGAAAGCAAGCGGAGCGGAGCGCAAGATG 221
QY 181 cggagcagcagaagagcagcggagacagaagaagcggcgaggagagagagc 240
Db 220 CGGAGACACGACAGAGAGACGACGGGAGCAAGAGAGCGGACCGCGGAGAGAGCG 161
QY 241 cgaagaagcggagagccgcgaggaagtgctcacatcaacgaacgatggaagaagac 300
Db 160 CGCAAGAGCGCGGAGGCCCGCAGGAGTGTCTGCATCACCGAAGCATGAGAGAG 101
QY 301 tacagcgacaaagtgaagccagccactggaagtcgacgcccgcctcgccgcggagag 360
Db 100 TACAGCGACAAAGTGAAGCCAGCCACTGAGTGCAGCCCGCTCGCGCCGCGGAG 41
QY 361 cggctcgaattggagagcggagacagtaagaagaagaatggaagaagaa 400
Db 40 CGGTTCCAGTTGGGAGACGGCGGAGCCAGTAAATAAAG 1
```

RESULT 10

BG574409 804 bp mRNA EST 10-APR-2001

LOCUS 602596352p1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:4705189 5',

DEFINITION mRNA sequence.

ACCESSION BG574409

VERSION BG574409.1 GI:13582062

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 804)

AUTHORS NIH-MGC http://mgc.ncbi.nlm.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaps-r@mail.nih.gov

Tissue procurement: DCPD/DRP

CDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM10571 row: 9 column: 14

High quality sequence stop: 686.

FEATURES

source location/Qualifiers

1..804

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4705189"

[illegible]

Db	255	ACAGGAGAGCAAGAACCAGCTGGCCCGAGTCTCCTTCAGCGGAATCAGGCTCAGTGCTGTG	314
Qy	509	aggaagaaggaggaaggaaaggaaggcggaggaaggaaggacccagtgaaataatca-	567
Db	315	AGGAAGAAGAGAGAGAGAGGAAAAAGGAGGAGGAGAGGAAGGAGGCACGACTCAAAATAACG	374
Qy	568	-----gaggaagaagaaggagggaggaaggaaccgcgagcaccttgaagg	619
Db	375	AGGAGAGAGAGAGAGAAAGAGGAAGGAGGAGAGGAGACGCCGACACTCTGAGGAGG	434
Qy	620	catacagagcagtctgcggaagtaagttgaagaaatat	660
Db	435	CATCAGACAGACTCTCCGAAGAAGTAAGTAGAGCACGAATG	475
RESULT	8		
LOCUS	BC009375	906 bp mRNA	HNC 09-JUL-2001
DEFINITION	Homo sapiens, similar to cell division cycle 2-like 1 (PITSLRE proteins), clone IMAGE:4121554, mRNA.		
ACCESSION	BC009375		
VERSION	BC009375.1	GI:14627288	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 906) Strausberg,R. Direct Submission Submitted (12-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA NIH-MGC Project URL: http://mgc.nci.nih.gov		
REFERENCE	Contact: MGC help desk Email: cgapbs-ref@mail.nih.gov		
AUTHORS	Tissue Procurement: ATCC		
JOURNAL	CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland, Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@nhlbi.nih.gov Shenchenko,Y., Wetherby,K.D., Beckstrom-Sternberg,S.M., Benjamin,B., Blakeley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S., Dietrich,N.L., Guan,X., Gupta,J., Ho,S.-L., Karlins,E., Legaspi,R., Lim,M., Maduro,Q.L., Mastello,C., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Snyder,B., Stantipop,S., Thomas,P.D., Zhang,L.-H. and Green,E.D.		
REMARK			
COMMENT			
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL plate: 26 Row: d Column: 15 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 3850303 This clone has the following problem: incomplete processing.		
FEATURES			
source	Location/Qualifiers 1..906 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4121554" /tissue_type="Muscle,"Thadomyosarcoma" /clone_id="NH_MGC_17" /lab_host="DH10B-R" /note="Vector: pOTB7"		
BASE COUNT	317 a 170 c 307 g 112 t		
ORIGIN			
Query Match	51.0%; Score 336.6; DB 12; Length 906;		

Query Match 71.9%; Score 474.8; DB 10; Length 1042;
 Best Local Similarity 89.0%; Pred. No. 1.4e-68;
 Matches 596; Conservative 0; Mismatches 32; Indels 42; Gaps 6;

QY 5 aaagtcggaagacgacgtctgaagaaacagataaagctcgccgggaatggaagacaga 64
 |||||||
 Db 1 AAGGTGCGAAACCGCATCAGAAAGAACAGATTAACCTCGCGGAAATGGAAAGACAGA 60
 65 agagaagggaaatggaagag-aaagcattccagagagaaaggggaatgatgcygtgc 123
 |||||||
 Db 61 AGGAAGAGGAGATGGCAAGGTGAGCTTCACAGAGAGAA----- 99
 QY 124 ctcttcagagaccgcttggagcagttagaagaagcggagacggagacgcaagatcg 183
 |||||||
 Db 100 -----AGGAGCCGCTTGGAGCAGTTAGAAAGAGCCGGAGGCGGAGCAAGATCGCG 153
 QY 184 gagcagcagaagagcagcgggagcagaagagcgcgagcggcgagagagcgagcgc 243
 |||||||
 Db 154 GACACACAGAGAGAGAGCGGGGAGCAAGAGAGCCGAGCGGGGAGAGAGCGCGCC 213
 QY 244 aagagcggagagcccgacggagagtgctgcacatcacgagacgattgagagagctac 303
 |||||||
 Db 214 AAGGAGCGGAGGCGCCGACAGGAAAGTGTCTGCACATCACGAACGATGAGAGGACTAC 273
 QY 304 agcgacaaagtgaagccagccacttgagtcgagccgcctcgagccgcgcgagagcg 363
 |||||||
 Db 274 AGCGACAAAGTGAAGCCAGCCACTGAGTCCGACGCCGCTCGCGCGCGGAGCGG 333
 QY 364 ttgagttggagaaagcggcggaagccagtaa----aagaagagaaatggaagagga 419
 |||||||
 Db 334 TTGAGTGGGAGACGGCGGAGCCAGTAAACGAAGAAACTGCGCAAGCAAGAGGGA 393
 QY 420 cctgcgtccgacttcagagacatcagcagcagcagagaaagacccagctcgccgagtc 479
 |||||||
 Db 394 CCGGCTGTCGATTAAGAGACATCAAGCAGCAGAGAGAAAGACCAAGCTCGCGCAGTC 453
 QY 480 ctgcagcagaatcaggtcaggtctcgaagaaagagagagagagagag--agag 537
 |||||||
 Db 454 CTGCTGACGCGAATCAGGCTCAGGTTCTGAGGAAGAGAGAGAGAGCAAGCAAGAG 513
 QY 538 gaagagagaaagagcaccagttgaagaaatcagagagagagaaagaaagagagagga 596
 |||||||
 Db 514 ACGGAGAGCGGAGACACCACTGAAGATCAGAGAGAGAGAGAGAGAGAGAGAGAG 573
 QY 597 -----gaccggcagcaactctgaagagcatcagagcagtcgcccgaagaaagtgta 650
 |||||||
 Db 574 GAGGAGGAGACGAGAGAACTCTGAGGAGGCAATCAGAGCAAGTCTG-CGAAGAACTAAGTGA 632
 QY 651 ggaagaaatg 660
 |||||||
 Db 633 GGAAGAAATG 642

RESULT 5
 LOCUS BG822824 662 bp mRNA EST 22-MAY-2001
 DEFINITION 602727816F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4867443 5',
 mRNA sequence.
 ACCESSION BG822824
 VERSION BG822824.1 GI:14170411
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 662)
 AUTHORS NIH-MGC <http://mgc.mci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-f@mail.nih.gov

Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: NIH Intramural Sequencing Center
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.nih.gov>
 Plate: L10CM1735 row: p column: 04
 High quality sequence stop: 601.
 Location/Qualifiers
 1. 662

FEATURES

source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4867443"
 /clone_1db="NIH_MGC_15"
 /tissue_type="adenocarcinoma cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: colon; Vector: pOM7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCAAGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)"
 BASE COUNT 202 a 135 c 255 g 66 t
 ORIGIN

Query Match 63.6%; Score 420; DB 11; Length 662;
 Best Local Similarity 99.5%; Pred. No. 1.3e-55;
 Matches 442; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 188 agcagaagagacagcgggagcagaagagcgcgagcggcgagagagcgccgaag 247
 |||||||
 Db 2 AGCAAGAGAGAGAGCGGAGCAGAAAGAGCGCGAGCGCGGAGAGCGGCAAG 61
 QY 248 agcggagagcccgcaaggatgtctgcacatcacgaagatgagagagactaacgcg 307
 |||||||
 Db 62 AGCGGAGGCGCGGAGGAAAGTGTCTGCACATCACGAAAGATGAGAGAGACTACAGCG 121
 QY 308 acaagtgaaagccagccagctgagatgcagcccgccctgcgcgcggagagcggttcg 367
 |||||||
 Db 122 ACAAGTGAAGCCAGCCACTGAGATCGACGCCCTCGGCGCGGAGAGCGGTTCCG 181
 QY 368 agttggagagcggccggagagcagtaaaagaaatggaagaaagggacctgtgt 427
 |||||||
 Db 182 AGTTGGAGAGCGCGCGGAGGACAGTAAGAAAGAAATGGAAGAAAGGAACTGCTGT 241
 QY 428 ccgacttacaggaacatcagcagcagcagagaaagacagctcgccgagctcctcag 487
 |||||||
 Db 242 CCGACTTACAGGACATCAGCAGACCGAGAGAGACCAAGTCCGCGAGTCTGTCAG 301
 QY 488 cagaatcaggtcaggtctcgaagaaagagagagagagagagagagagagagag 547
 |||||||
 Db 302 CAGAAATCAGGCTCAGGTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
 QY 548 ggaagcacc-agtgaagaaatca-gagagagagagagagagagagagagagagag 605
 |||||||
 Db 362 GGAGCAGCGAGTGAAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 421
 QY 606 caactctgagagagcatcagagca 629
 |||||||
 Db 422 CAACCTGAGGAGGATCAGAGCA 445

RESULT 6
 LOCUS AK008283 1000 bp mRNA HTC 05-JUL-2001
 DEFINITION Mus musculus adult male small intestine cDNA, RIKEN full-length
 enriched library, clone:2010016J05, full insert sequence.
 ACCESSION AK008283
 VERSION AK008283.1 GI:12842370

OY	486	aggaaatcgaagcttaagtttcgaagaagaa-----gaagaaaggaaggaagaaggaag	541
Db	454	ACCGGAATCAGGCTTAGTTCAGGAGAACAACGAGAGAGGCCGACAGACAGAGAAC	513
OY	542	aggaaaggagaccacagttaaagatca-----gaagaaagaaagaaaggaaggaag	592
Db	514	GGACAGGGAGAGCCACGTGACCAATCATCAGAGAGAGAGAGAGAGAAAGAGAGAGAGG	573
OY	593	aggaaaccggcgacgaactctgaagagaca-tcagaagcagttccgaaagaagtagag	651
Db	574	AGGAACACCGCACGCAACTCTGAGGAGGCATTTCAGAGCATGTGCGCAAGAGTAAGTGAG	633
OY	652	gaagaaatcy 660 	
Db	634	GAGAGAAATG 642	
RESULT	3		
BGI07702			
LOCUS	BGI07702	1340 bp	mRNA
DEFINITION	60227878F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4365474 5'	EST	30-JAN-2001
ACCESSION	BGI07702		
VERSION	BGI07702.1	GI:12601548	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 1340)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LHAM10014 row: 1 column: 19 High quality sequence stop: 667. Location/Qualifiers 1. /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:4365474" /clone_id="NIH_MGC_86" /tissue_type="osteosarcoma, cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: bone; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.53 kb. Library enriched for full-length clones and constructed by Life Technologies Notes: This is a NIH-MGC Library."		
FEATURES	source		
BASE COUNT	467 a	272 c	450 g 151 t
ORIGIN			

Query Match	Similarity	Score	4/6.4	DB 11	Length 1340
Best Local	Similarity 98.0%	Pred. No. 7.7e-69			
Matches 496	Conservative 0	Mismatches 1	Indels 9	Gaps 1	
OY	164	agcggagcgcagaatctgaggacagcagaagcagcgagcagcagaagcagaagcagcagc	223		
Db	1	AGCGGAGCGCAAGATCGCGGAGCGAGCAGCAAGAGCAGCAGCGGAGCGAGCGGAGCG	60		
OY	224	ggcggcgaggagagagcgcaaggagcgggagcccgacgggaagtgtctgcaatcacc	283		
Db	61	GCGGCGCAGAGAGCGGCGCAAGAGCGGAGACCCCGCAGGGAAGTGTCTGCAATCACC	120		

QY	284	gacagatgagagagactacacgacgaatgtaagccagccactgtgagtcgagccgc	343
Db	121	GAACGATGAGAGGACTACACGCAAAAGTGAAGCCAGCCACTGGAGTGCAGCCCGC	180
QY	344	cctgcgcgcgcgcgcgcgcgtttcagttctgttggaagacgcgcgcgaagccagtaaaagaaga	403
Db	181	CTCGGCGCGCGGGGAGCGGTTCCAGTTTGGGAGACGGCCGGAAGCCAGTAAAGAAAGAGA	240
QY	404	aaatggaagaaaggagaccctgtctcgcacttaccaggaatcaagcgaagcgaagagaaga	463
Db	241	AAATGAGAAAGAGGAGCTGCTGCTCCGACTTACAGGACATCAGCGACAGCGAGGAAGA	300
QY	464	ccagctgcgcgcgcgcgcgcgtctcgttcacagaatacagctcaggttcttgaggaagaagaag	523
Db	301	CCAGCTCGCGCCGAGTCCCTGCTCAGCGGAAATCAGGCTCAGGTTCTGAGGAAGAGGAGG	360
QY	524	aggagaagagaagagaagagaagagaaggaacccacagtgaaatca-----gaggaag	574
Db	361	AGGAGAGAGAGAGAGAGAGAGAGAGGAGACCCAGTGAAGATCTCAGAGAGAGAGAGAGAGG	420
QY	575	aagagaagagaagagaagagaagagaccgcgaacactcaggaagagcatcagaagcagctc	634
Db	421	AAGAGAGAGAGAGAGAGAGAGAGAGGAGACCGGCACACACTCTGAGAGAGCATCAGAGCAGTCTG	480
QY	635	ccgaagaagtaagtgaaggaagaatg	660
Db	481	CCGAAGAAGTAAGTGCAGAGAGGAATG	506
RESULT	4		
LOCUS	BE742943	1042 bp	mRNA
DEFINITION	601574925f1 NIH_MGC_9	Homo sapiens	CDNA clone IMAGE:3835686 5',
ACCESSION	BE742943		
VERSION	BE742943.1	GI:10157033	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 1042)		
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: DCTD/DTF cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov Plate: LCM518 row: b column: 07 High quality sequence stop: 667. Location/Qualifiers 1. 1042 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3835686" /clone_id="NIH_MGC_9" /tissue_type="adenocarcinoma cell line" /lab_host="DH10B (phage-resistant)" /note="Organ: ovary; Vector: pOMB7; Site1: XhoI; Site2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(g). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."		
FEATURES	source		
BASE COUNT	371 a	159 c	390 g
ORIGIN	122 t		

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:14:23 : Search time 3881.49 Seconds
(without alignments)
1827.188 Million cell updates/sec

Title: US-09-915-060-6

Perfect score: 660

Sequence: 1 caggaacgtcgaacgaca.....aagtaagtgcgagaagaatg 660

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estln:*
4: em_estom:*
5: em_estpl:*
6: em_estda:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: gb_estli:*
11: gb_estc2:*
12: gb_hic:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_loy:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rtd:*
20: em_gss_vtl:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	568.2	86.1	1063	10	BE740559 601595656
2	518.2	78.5	723	10	BE742222 601575629
3	476.4	72.2	1340	11	BG107702 602277878
4	474.8	71.9	1042	10	BE742943 601574925
5	420	63.6	662	11	BG822824 602727816
6	362.8	55.0	1000	12	AK008283 60156008
7	339.4	51.4	516	11	BF156008 RC0-H7095
8	336.6	50.4	906	12	BC009375 Mus muscu
9	332.8	50.4	592	10	AT580044
10	314.2	47.6	804	11	BG574409 tq45g08.x
11	310.6	47.1	493	11	BI018825 IL3-MT026
12	286.4	43.4	830	11	BF607177 MY1_00011

13	277.6	42.1	812	11	BT149024 602911290
14	271.6	41.2	326	11	BG943585
15	266	40.3	758	11	BG969359
16	265.8	40.3	826	11	BG034678
17	256.6	38.9	372	10	AL034747
18	244.4	37.0	740	10	AL043340
19	212.6	32.2	367	10	AA681955
20	208.2	31.5	345	11	BT261348
21	205.6	31.2	674	10	BE259507
22	170.8	25.9	649	10	AW210026
23	167.2	25.3	283	10	AW804794
24	166.8	25.3	251	10	AV270235
25	166.6	25.2	476	10	BE680738
26	164.4	24.9	900	11	BG298479
27	155.6	23.6	1083	11	BG770237
28	153	23.2	537	11	BF743770
29	149.2	22.6	253	10	AV327502
30	142.2	21.5	292	10	AV368285
31	138	20.9	441	10	AW451883
32	135.8	20.6	233	10	BR290670
33	135.6	20.5	182	10	AA413247
34	134.2	20.3	456	11	BF998861
35	127.6	19.3	725	11	BF308197
36	126.4	19.2	216	10	AV105652
37	124	18.8	514	10	AI989489
38	120.8	18.3	462	11	BG486162
39	119.4	18.1	732	10	AI006623
40	110.2	16.7	160	10	AM193827
41	107.6	16.3	1005	10	AL523189
42	106	16.1	673	11	BF983533
43	104.4	15.8	469	11	BF359805
44	104	15.8	339	11	H33616
45	103.8	15.7	321	10	AA309711

ALIGNMENTS

RESULT 1
LOCUS BE740559 1063 bp mRNA
DEFINITION 601595656F1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3949693 5',
mRNA sequence.
ACCESSION BE740559
VERSION BE740559.1 GI:10154551
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1063)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DMP

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMU at: image.llnl.gov
Plate: L1CMB14 row: p column: 14
High quality sequence stop: 741.
Location/Qualifiers

FEATURES
source

1..1063
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3949693"
/clone_lib="NIH_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOTB7; Site_1: XhoI; Site_2:

Qy 628 cagtcctgcgaaagaagtaagtgaagaaga 656
| | | | | | | | | | | | | | | | | |
Db 106 GGGCAGGAGGAGGAGGAGGAGGAGGAGGA 78

Search completed: December 4, 2001, 18:58:37
Job time: 9263 sec

Matches	241: Conservative	0: Mismatches	271: Indels	0: Gaps
Qy	141	ggagcagcttagaagaagacgagcgagcgaggagcgaagatgcgaggagcaagaagaagagca	200	
Dy	141			
Dy	949	ggggcagagacagacagagagggcgacagagggcgacagacagacagagggcgacagagggcgca	1008	
Qy	201	ggcgagagcagaagagcgcgagcgcgaggcgaggagcgcgacagagagcgagagagcgagcgccg	260	
Dy	1009	ggagggcgacagacagacagagagggcgacagacagacagagggcgacagagggcgacagagggcgca	1068	
Dy	1009			
Qy	261	cagggaagtgtctgcacatccacgaaacgatgagaaagactacagcgacaagvtgaagc	320	
Dy	1069	ggagcagagagggcgacagacagacagagagggcgacagagggcgacagagggcgacagagggcgca	1128	
Qy	321	caagccctgagtgcgagcgcccgcccgcccgaggagaggttcagatgttgagaacg	380	
Dy	1129	ggggcagagacagacagagggcgacagacagacagagggcgacagagggcgacagagggcgca	1188	
Qy	381	ccggaagccagtaaaagaagaataatgaaagaagagaccgtgtctgcgacttacagga	440	
Dy	1189	ggagcagagagggcgacagagagggcgacagacagagagggcgacagagggcgacagagggcgca	1248	
Qy	441	calcagcgacagcgagaggaagacacagctcgcgagatccctgcagcagatcagctc	500	
Dy	1249	ggggcagagagggcgacagacagacagagagggcgacagagggcgacagagggcgacagagggcgca	1308	
Qy	501	aggttcaggaagaagaagaggaagaagaagaagaagaagaagaagaagaagaagaagaaga	560	
Dy	1309	ggggcagagagggcgacagacagagggcgacagagggcgacagagggcgacagagggcgacagagggcgca	1368	
Qy	561	agaatcagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga	620	
Dy	1369	ggagagagggcgacagacagacagagggcgacagagggcgacagagggcgacagagggcgacagagggcgca	1428	
Qy	621	atcagagcagctgtccgaagaagtgaagtgaag	652	
Dy	1429			
Dy	1429	agtgagagccgggggtcgagagagctagtggagc	1460	
RESULT	6			
US-07-884-811-15				
Sequence 15, Application US/07884811				
Patent No. 5316921				
GENERAL INFORMATION:				
APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.				
TITLE OF INVENTION: SINGLE-CHAIN HEPATOCYTE GROWTH FACTOR VARIANTS				
NUMBER OF SEQUENCES: 21				
CORRESPONDENCE ADDRESS:				
ADDRESSEE: Genentech, Inc.				
STREET: 460 Point San Bruno Blvd				
CITY: South San Francisco				
STATE: California				
COUNTRY: USA				
ZIP: 94080				
COMPUTER READABLE FORM:				
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk				
COMPUTER: IBM PC compatible				
OPERATING SYSTEM: PC-DOS/MS-DOS				
SOFTWARE: paln (genentech)				
CURRENT APPLICATION DATA:				
APPLICATION NUMBER: US/07/884, 811				
FILING DATE: 19920518				
CLASSIFICATION: 530				
PRIOR APPLICATION DATA:				
APPLICATION NUMBER:				
FILING DATE:				
ATTORNEY/AGENT INFORMATION:				
NAME: Dreger, Ginger R.				
REGISTRATION NUMBER: 33, 055				
REFERENCE/DOCKET NUMBER: 755.1				
TELECOMMUNICATION INFORMATION:				
TELEPHONE: 415/225-3216				
TELEFAX: 415/952-9881				

```

      / TELTEX : 910/371-7168
      / INFORMATION FOR SEQ ID NO.: 15:
      / SEQUENCE CHARACTERISTICS:
      / LENGTH: 10596 bases
      / TYPE: NUCLEIC ACID
      / STRANDEDNESS: single
      / TOPOLOGY: linear
      / US-07-884-811-15

Query Match          11.9%; Score 78.4; DB 1; Length 10596;
Best Local Similarity 47.1%; Pred. No. 6e-08;
Matches 241; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY   141 ggaagcaqttaagaagaagcgcgagcgcgagacgaagtatcggyagccagacaagagaga 200
     ||| |||| |
Db    2483 GGCGCAGACGACGACGAGAGGGCCAGAGGGGCAGAGGCAAGAGAGGGCCAGAGGGGCA 2542

QY   201 gcggaggacagaagaagcgcgagcgcgagcgcgagagagcgagcgacaagagacgagccccg 260
     ||| |||| |
Db    2543 GGAGGGGCCAGAGAGCAGAGAGAGGGGCGAGGACGAGGCGGCGAGAGGGCCAGAGGGGCA 2602

QY   261 cagggaaagtgtcttcacatccaccgaagcatgagagagagagactaacgcacaagtgaagc 320
     ||| |||| |
Db    2603 GGAGCAGAGAGGGGCGACAGAGAGAGGGGCGAGGCGGCGAGAGGGCGAGAGCAGAGA 2662

QY   321 cagccactgtagtgcagaccgcccgctcgccgcgcgcgcgagcgagtgltcgagttgagagcyg 380
     ||| |||| |
Db    2663 GGCGCAGAGCAGAGAGAGGGCCAGAGCAGGAGAGGGCGCACAGAGAGGGCCAGAGGGGCA 2722

QY   381 ccggaagccagttaaagaagaagaataatgaaagaagaaggaacctgtgtcttcgacttacaga 440
     ||| |||| |
Db    2723 GGAGCAGAGAGGGGCGCAGAGGGCGCAGAGCAGAGAGGGGCGAGAGGGGCGAGCAGAGA 2782

QY   441 catcagcgacagcgagaggaagaaccaagctcgagtcgcgagtcctcgtcaagcagaatcagctc 500
     ||| |||| |
Db    2783 GGCGCAGAGGGGCGACAGAGCAGAGAGGGGCGACAGAGGGGCGACAGAGGGGCGAGAGA 2842

QY   501 aggttcaggaaagaagaagagagagagagaaagaagaagaagaagaagagcaccaatga 560
     ||| |||| |
Db    2843 GGCGCAGAGCAGAGAGGGGCGACAGAGGGGCGACAGAGCAGAGGGGCGAGAGGGGCGACA 2902

QY   561 aaaaacagaagagagagagagagagagagagagagagagacccgcagcaactctgagagagc 620
     ||| |||| |
Db    2903 GGAGCAGGGGCGCAGAGCAGAGGGGCGAGGAGGAGGTGAAGCCGGGCTCGAGAGAGGC 2962

QY   621 atcagagcagctctgccgaagaaglaagtaag 652
     ||| |||| |
Db    2963 ACTGGAGGCCGGGCTCGAGAGGTAOTGACAG 2994

RESULT
US-07-885-971-15
; Sequence 15, Application US/07885971
; Patent No. 5328837
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J., Lokker, Nathalie A., Mark, Melanie R.
; TITLE OF INVENTION: HEPATOCYTE GROWTH FACTOR PROTEASE DOMAIN VARIANTS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/885,971

```

Matches 241; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

```
OY 141 ggagcagttagaagaagcggagcggagcgaagatgcggagagcagcagaaggaaca 200
    || |||||
DB 903 GGGGCGAGGAGCAGAGAGAGGGGCGAGAGGCGCAGAGAGAGAGAGGGGCGAGAGGGGCA 962
OY 201 gcggagagcagaagagcgcgagcggcggagcggagcggagcggagcggagcggagcgg 260
    || |||||
DB 963 GGAGGGGCGAGGAGAGAGAGGGGCGAGAGGCGGCGAGAGGCGGCGAGAGGGGCA 1022
OY 261 cagggaagtgtctgcacatccccaagcagatgagaaagactacagcgcaaatgtaagc 320
    || |||||
DB 1023 GGAGGAGAGGGGCGAGGAGCAGAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGCAGA 1082
OY 321 cagccacgtgagtcgcagccgcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 380
    || |||||
DB 1083 GGGGCGAGGAGCAGAGAGGGGCGAGAGCAGAGAGGGGCGAGAGGGGCGAGAGGGGCA 1142
OY 381 ccggaagcagtaaaagaagaagaatgaaagaagagcctgtctgcacttacagga 440
    || |||||
DB 1143 GGAGCAGAGGGGCGAGAGGGGCGAGAGCAGAGAGGGGCGAGAGGGGCGAGAGAGGA 1202
OY 441 catcagcgacagcgagaggaagaccagctcgcgagtcctcgtcagcagaatcagctc 500
    || |||||
DB 1203 GGGGCGAGAGGGGCGAGAGGAGCAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGCA 1262
OY 501 aggtctgaggaagaagaagaagaagaagaagaagaagaagaagaagaagaagcagcagga 560
    || |||||
DB 1263 GGGGCGAGAGGCGAGAGGGGCGAGAGGGGCGAGAGGCGCGAGAGGGGCGAGAGCA 1322
OY 561 agaatcagaggaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagc 620
    || |||||
DB 1323 GGAGAGGGGCGAGAGGAGCAGAGGGGCGAGAGGCGAGAGGCGCGGCGCGAGAGGAGC 1382
OY 621 atcagaagcagctcgcggaagaagtaagtgaag 652
    || |||||
DB 1383 AGTGGAGGCGGGGCGTGTGAGAGTAGTGAGG 1414
```

RESULT 4

```
US-09-130-114-1/c
; Sequence 1, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; FILE REFERENCE: 0867/ID903US1
; CURRENT APPLICATION NUMBER: US/09/130.114
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5452
; TYPE: DNA
; ORGANISM: VEBNA
US-09-130-114-1
```

Query Match 11.9%; Score 78.4; DB 2; Length 5452;
Best Local Similarity 47.1%; Pred. No. 5.6e-08;
Matches 241; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

```
OY 141 ggagcagttagaagaagcggagcggagcgaagatgcggagagcagcagaaggaaca 200
    || |||||
DB 1902 GGGGCGAGGAGCAGAGAGGGGCGAGAGGCGCAGAGCAGAGAGGGGCGAGAGGGGCA 1843
OY 201 gcggagagcagaagagcgcgagcggcggagcggagcggagcggagcggagcggagcgg 260
    || |||||
DB 1842 GGAGGGGCGAGGAGGAGAGGGGCGAGAGGCGGCGAGAGGGGCGAGAGGGGCA 1783
```

```
OY 261 cagggaagtgtctgcacatccccaagcagatgagaaagactacagcgcaaatgtaagc 320
    || |||||
DB 1782 GGAGCAGAGAGGGGCGAGAGCAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGCAGGA 1723
OY 321 cagccactgtagtcgagccgcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 380
    || |||||
DB 1722 GGGGCGAGGAGCAGAGGGGCGAGAGGCGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCA 1663
OY 381 ccggaagcagtaaaagaagaagaatgagaaagagactcgtctgcacttacagga 440
    || |||||
DB 1662 GGAGCAGAGGGGCGAGAGGGGCGAGAGGCGGCGAGAGGGGCGAGAGGGGCGAGAGCA 1603
OY 441 catcagcgacagcgagaggaagaccagctcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 500
    || |||||
DB 1602 GGGGCGAGAGGGGCGAGAGCAGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCA 1543
OY 501 aggtctgaggaagaagaagaagaagaagaagaagaagaagaagaagaagaagcagcagga 560
    || |||||
DB 1542 GGGGCGAGAGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGCA 1483
OY 561 agaatcagaggaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagc 620
    || |||||
DB 1482 GGAGCAGGGGCGAGAGGAGGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGGGGCGAGAGG 1423
OY 621 atcagaagcagctcgcggaagaagtaagtgaag 652
    || |||||
DB 1422 AGTGGAGGCGGGGCGTGTGAGAGTAGTGAGG 1391
```

RESULT 5

```
US-08-910-647-1
; Sequence 1, Application US/08910647
; Patent No. 6251433
; GENERAL INFORMATION:
; APPLICANT: Zuckermann et al.
; TITLE OF INVENTION: Compositions and Methods for
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910,647
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Fujita, Sharon M.
; REGISTRATION NUMBER: 38,459
; REFERENCE/DOCKET NUMBER: 1218.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-2706
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9600 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-910-647-1
```

Query Match 11.9%; Score 78.4; DB 4; Length 9600;
Best Local Similarity 47.1%; Pred. No. 5.9e-08;

Query Match 15.1%; Score 99.4; DB 1; Length 7218;
 Best Local Similarity 5.4%; Pred. No. 3e-12;
 Matches 22; Conservative 259; Mismatches 130; Indels 0; Gaps 0;

```

QY 2 acgaagctcgaaacgacatcgagaaagaaacagataagctgcgcggaatggtgaaagac 61
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1469 ATGCAAGTAGTTAAAGATTAAGAAATTGGTACRRRRRRRRRRRRRRRRRRRRRRRRR 1410

QY 62 agaaagaaaggaatgcaagagagcatccagagaaagaggaaggaatggtggtcgt 121
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1409 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1350

QY 122 gcccttcagagcgctcttgagcagltagaagaaagcgagagcgagagcgcaaatgc 181
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1349 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1290

QY 182 gggagcagcagaagaaagcagcagagagagagagcgcgagcgagcgagagagcggc 241
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1289 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1230

QY 242 gcaagagcgagagcgccgagagagagtgctgcacatccagacgatgagagagact 301
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1229 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1170

QY 302 acagcgcaaaagtgaaagcgacccactgagtgctgcagccgctcgcgcgagcgagagc 361
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1169 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1110

QY 362 ggttcagtgtaggagcgcgagcagcagcagtaaaagaaagaaatggaag 412
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1109 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1059
  
```

RESULT 2

```

US-09-130-114-2/c
; Sequence 2, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Dana, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; FILE REFERENCE: 0867/ID9030US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: EBNA
US-09-130-114-2
  
```

Query Match 12.8%; Score 84.8; DB 2; Length 1931;
 Best Local Similarity 46.6%; Pred. No. 2.5e-09;
 Matches 272; Conservative 0; Mismatches 312; Indels 0; Gaps 0;

```

QY 72 ggaatgcaagggagcatctcagagagaaagggaaatgtagtgcgtctcttcag 131
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 953 GGACGGGAGGAGAGACGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 894

QY 132 ggaacgcttgagcagcttgaaagagcgagagcgagagcgcaagatgagagagcaga 191
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 893 GGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 834

QY 192 gaagagcagcgagagcagaagagcgagcgagcgagagcgagcgagcgagcgagcg 251
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 833 CGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 774

QY 252 ggaagcccgagagagtgctgcacatccgaacgatgagagagactacagcgacaa 311
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
  
```

```

Db 773 CGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 714
QY 312 agtgaagccagaccactgtagtgcagccgctcgcccgcgagcgagtgtagt 371
Db 713 GGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 654
QY 372 ggaagacggcggaagccagtaaaagaaagaaatgtaagaaagagccgtctccga 431
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 653 CGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 594
QY 432 cttaacagacatcagcagcagcagagaaagacacagctcgccagatctctgcagcaga 491
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 593 CGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 534
QY 492 atcagctcaggttcttgaaagaaagagagagagagagagagagagagagagag 551
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 533 GGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 474
QY 552 caacagtgagatcagaagagagagagagagagagagagagagagagagagagagc 611
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 473 GGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 414
QY 612 tgaagagcattagagcagtgctgcgaagaaatgagtgaggaag 655
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 413 GGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGC 370
  
```

RESULT 3

```

US-09-050-863-2
; Sequence 2, Application US/09050863
; Patent No. 6114111
; GENERAL INFORMATION:
; APPLICANT: Lao, Ying
; APPLICANT: Hwang, Betty
; APPLICANT: Payan, Don
; TITLE OF INVENTION: Mammalian Protein Interaction Cloning
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESS: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/050,863
; FILING DATE: 30-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silva, Robin M.
; REGISTRATION NUMBER: 38,304
; REFERENCE/DOCKET NUMBER: A-65638/DJB/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 949-8711
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2580 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA
US-09-050-863-2
  
```

Query Match 11.9%; Score 78.4; DB 3; Length 2580;
 Best Local Similarity 47.1%; Pred. No. 5.1e-08;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:57:53 ; Search time 152.61 Seconds
(without alignments)
979.461 Million cell updates/sec

Title: US-09-915-060-6

Perfect score: 660

Sequence: 1 caccagacgtcggaacagaca.....aagtaagtggagaaagtg 660

Scoring table:

IDENTITY_NUC
Gapox 10.0 , Gapext 1.0

Searched: 351203 segs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfilest1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	99.4	15.1	7218	1 US-08-232-463-14	Sequence 14, Appl
C 2	84.8	12.8	1931	2 US-09-130-114-2	Sequence 2, Appl
C 3	78.4	11.9	2580	3 US-09-050-863-2	Sequence 2, Appl
C 4	78.4	11.9	5452	2 US-09-130-114-1	Sequence 1, Appl
C 5	78.4	11.9	9600	4 US-08-910-647-1	Sequence 1, Appl
C 6	78.4	11.9	10596	1 US-07-884-811-15	Sequence 15, Appl
C 7	78.4	11.9	10596	1 US-07-885-971-15	Sequence 15, Appl
C 8	78.4	11.9	10596	1 US-08-087-783A-15	Sequence 15, Appl
C 9	78.4	11.9	10596	1 US-08-194-088B-15	Sequence 15, Appl
C 10	78.4	11.9	10596	2 US-08-194-088B-15	Sequence 15, Appl
C 11	78.4	11.9	10596	5 PCT-US93-04648-15	Sequence 15, Appl
C 12	75.4	11.4	3489	2 US-08-728-323A-1	Sequence 1, Appl
C 13	75.4	11.4	32207	4 US-08-770-379-20	Sequence 20, Appl
C 14	75.4	11.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 15	75.4	11.4	32207	4 US-08-781-891-209	Sequence 209, Appl
C 16	62.6	9.5	53526	3 US-08-658-136-1	Sequence 1, Appl
C 17	62.6	9.5	53526	3 US-08-658-136-1	Sequence 1, Appl
C 18	60.6	9.2	3211	2 US-08-574-959A-8	Sequence 8, Appl
C 19	60.6	9.2	3211	4 US-09-357-014-8	Sequence 8, Appl
C 20	60.6	9.2	3901	2 US-08-574-959A-6	Sequence 6, Appl
C 21	60.6	9.2	3901	4 US-09-357-014-6	Sequence 6, Appl
C 22	60.4	9.2	9551	1 US-08-056-200-93	Sequence 93, Appl
C 23	60.4	9.2	9551	2 US-08-800-644-93	Sequence 93, Appl
C 24	58	8.8	12001	1 US-08-458-568A-11	Sequence 11, Appl
C 25	57.8	8.8	15376	3 US-08-785-420-1	Sequence 1, Appl
C 26	57.8	8.8	43795	3 US-08-742-185-101	Sequence 101, Appl
C 27	57.6	8.7	2150	2 US-08-861-464-13	Sequence 13, Appl

C 28	57.6	8.7	2150	2 US-08-396-001-13	Sequence 13, Appl
C 29	57.6	8.7	2150	4 US-09-323-433A-13	Sequence 13, Appl
C 30	56.4	8.5	2340	3 US-09-022-983-4	Sequence 4, Appl
C 31	56.4	8.5	2477	4 US-09-490-692-3	Sequence 3, Appl
C 32	56.2	8.5	9636	1 US-08-323-170B-1	Sequence 1, Appl
C 33	55.6	8.4	16442	3 US-08-781-891-208	Sequence 208, Appl
C 34	54.8	8.3	397	3 US-09-253-691-3	Sequence 2, Appl
C 35	54.2	8.2	2277	1 US-08-676-967-2	Sequence 2, Appl
C 36	54.2	8.2	2277	1 US-08-676-974-2	Sequence 2, Appl
C 37	54.2	8.2	2277	2 US-09-098-487-2	Sequence 2, Appl
C 38	53.8	8.2	2301	1 US-08-306-691B-23	Sequence 23, Appl
C 39	53.8	8.2	2301	5 PCT-US93-06251-78	Sequence 78, Appl
C 40	53.6	8.1	19056	4 US-09-272-032-8	Sequence 8, Appl
C 41	53.2	8.1	4897	6 US-09-007-005-17	Sequence 17, Appl
C 42	53	8.0	289	4 US-09-244-796-17	Sequence 17, Appl
C 43	53	8.0	289	4 US-09-177-325-2	Sequence 2, Appl
C 44	53	8.0	1276	4 US-09-411-812A-2	Sequence 2, Appl
C 45	53	8.0	1276	4 US-09-411-812A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEFFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESS: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22113-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZ9pt-Fls
US-08-232-463-14

CC canine RPCR cDNA.

XX
SQ Sequence 2800 BP; 1026 A; 287 C; 938 G; 549 T; 0 other;

```
Query Match          12.1%; Score 79.6; DB 22; Length 2800;
Best Local Similarity 47.0%; Pred. No. 1.3e-08;
Matches 287; Conservative 0; Mismatches 314; Indels 9; Gaps 1;

QY 48 ggaatggnaaagacagagaaagggaaatgacaaagggaatccagagagaaaggg 107
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 ggaagaagggaattggaagaggaaggggaatggaatgaaggaaggaaggaaggg 600

QY 108 gaatgatgagctgtgctcttcacggccgcttgagcaattagaaggaagcgaggcg 167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 gaagaagaagaaaggggaattggaagggaagaaagagaggggaaggggaaggagg 660

QY 168 ggaacgcaagaatgagagagcagcagaaggaagcagcgagcagaagaagagcgagcg 227
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 661 aaagggagagctggaagggagaggaaggaagcagaaatggaaggaaggaagggagtgag 720

QY 228 ggcggagagcgagcgcaagggagcgccgcaaggaagtgctgcacatcacgcgaac 287
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 721 gaaggaagaaggggaagtggaaggaaggaaggaagtggaaggaaggaaggaagga 780

QY 288 gatgaagagagactacacgacaaagtgaagccaagccactggaatgctgcagccgctcg 347
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 781 ggaagagggggaagtggaag-----gaaggaagcagaagaagaagagatgcagaagaa 831

QY 348 gccgcgcgagagcgagctgagtgagagcagcgccggaagccagtaaaagaagaagaat 407
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 832 gaagaaaggggaagtgaagaagaagagaggggaaggaagaaataatgaaaggggagga 891

QY 408 ggaagaaaggggaacctgctgtccgacttacagggacatcagcgacagcgaaggaagaccag 467
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 892 ggaagaggggaagaaaggaagaaagaaatgaaagagagaggaagaaagagagagga 951

QY 468 ctgcgcgagagctcctcgtcagcagaatcaggtctcaggtcttgaggaagaagaagagagga 527
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 952 agcagggggaaggggaaggaaggaaggaaggggaagcagcagcaggaagggggggga 1011

QY 528 ggaagagagagaggaaggaaggaagcaccagtgaaagaaatcagagaggaaggaaggaagga 587
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1012 agaaagggggaagagagtgaaagggaaggaagggcgcaaggaaggaagaagaaggga 1071

QY 588 ggaagagagagagcggcagcgaactctgaggaagcattcagagcagtgctccgaagaagtgaag 647
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1072 ggaaggaaggaaggaaggaaggaaggggaaggaaggaaggaaggaaggaaggaagggag 1131

QY 648 tgaggaagaa 657
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1132 aggaagggaa 1141
```

Search completed: December 4, 2001, 19:03:47
Job time: 6988 sec

[illegible][illegible]

QY	201	gccggagcagaaagacgcgaagcgagcgagcgagagagccgcgaagaagacgcgagcccg	260
Db	471	ggagaggagaaagagggagagagaaagagaaagagagggtagagagagaaagagaggg	530
QY	261	cagggaagctgtctgcacctgcccaatctgaagagagatacacagcgacaagaagaaagc	320
Db	531	gtacagagagagaaagagagagagagagagaaagagaaagagagaaagaaagaaag	590
QY	321	cagccacatggagctccaaccccgcctcgaccgcgcgcgagagcgtctcgatttvggagacgg	380
Db	591	aagaaagagagaaagaaagaaagagatcatgataagagagagggagagagaaagag-gagga	649
QY	381	cccgagacccagctaaaagaagagaaaaatctgaagaagaaggactgtcttcgacttacsgga	440
Db	650	agagagaaagagagagagagagagaaagagaaagaaagagagagagagagagaaagga	709
QY	441	catcagcgacacagcgagagaaagacacagctcgcccgagtccctcgacagatatcagctc	500
Db	710	agagagagagagaaagagagagagagagagagaaagagaaagagaaagagagagagga	769
QY	501	aggtctctagagaaagagagagagagaaagagagagagagagagagagagagagagagagag	560
Db	770	ggagga	829
QY	561	agaaatcag	615
Db	830	gaaag	884
RESULT 11			
AAI35486			
ID	AAI35486 standard; DNA; 1969 BP.		
XX	AAI35486;		
AC			
XX	17-OCT-2001 (first entry)		
DE	Probe #4172 used to measure gene expression in human placenta sample.		
XX			
KW	Probe; microarray; human; placenta; antenatal diagnosis;		
XX	genetic disorder; ss.		
OS	Homo sapiens.		
PN	WO200157272-A2.		
PD	09-AUG-2001.		
XX			
PF	30-JAN-2001; 2001WO-US000663.		
XX			
PR	04-FEB-2000; 2000US-0180312.		
PR	26-MAY-2000; 2000US-0207456.		
PR	30-JUN-2000; 2000US-0608408.		
PR	03-AUG-2000; 2000US-0632366.		
PR	21-SEP-2000; 2000US-0234687.		
PR	27-SEP-2000; 2000US-0236359.		
PR	04-OCT-2000; 2000GB-0024263.		
XX			
PA	(MOLE-) MOLECULAR DYNAMICS INC.		
XX			
PI	Penn SG, Hanzel DK, Chen W, Rank DR;		
DR	WPI; 2001-488897/53.		
XX			
PT	Human genome-derived single exon nucleic acid probes useful for		
XX	analyzing gene expression in human placenta -		
PS	Claim 25; SEQ ID No 4172; 654bp; English.		
CC	The present invention relates to single exon nucleic acid probes (SENPs).		
CC	The present sequence is one such probe. The probes are useful for		
CC	producing a microarray for predicting, measuring and displaying gene		
CC	expression in samples derived from human placenta. The probes are useful		

Claim 25; SEQ ID No 4172; 654bp; English.

DR WPI; 2000-499331/44.
XX
PT Nucleic acids encoding internal ribosome entry sequences useful for
XX directing protein expression in gene therapy procedures -
XX
PS
XX Claim 4: Page 32; 57pp; English.
XX
CC The PITSURE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence is the PITSURE protein kinase IRES. The
CC IRES sequence and fragments of it may be used to induce cell cycle
CC dependent initiation of translation in eukaryotic cells. Vectors
CC containing the IRES may be used for the preparation of compositions for
CC the treatment of and/or prevention of a disease by gene therapy. Such
CC diseases may be cancer and restenosis.
XX
SQ Sequence 222 BP; 77 A; 36 C; 89 G; 20 T; 0 other;

Query Match	33.6%;	Score 222;	DB 21;	Length 222;
Post Local	100.00;			

```

best local similarity 100.0%;  pred. no. 4.4e-40;
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

[illegible][illegible][illegible][illegible]

Qy	619	gcacccagcagctctgcccgaagaagttaagtgcggagaatg	660
b	181	gcaccgaagcagctctgccgaagaatgaattgcgaacaata	222

RESULT 5

AAA73710
ID AAA73710 standard; RNA; 222 BP.
XX

AC	AAA73710;
XX	
DT	07-DEC-2000 (first entry)

XX RNA of human PIRLSRE protein kinase gene internal ribosome entry site.
XX

KM IRES; internal ribosome entry site; PITSARE; human; gene therapy;
KM cancer; restenosis; p58; p10; protein kinase; ss.
XX

Key	Location/Qualifiers
US	Homo sapiens.
XX	
FH	
XX	

```

FT      misc_signal      1.1222
FT      /tag= a
FT      /label= IRES
FT      /date= "1999-01-01"

```

FI	/note="Internal ribosome entry site"
XX	
PN	WO200044896-A1.
XX	

26-JAN-2000: 2000WG-EP00643	XX	PD	03-AUG-2000.
-----------------------------	----	----	--------------

XX 26 JAN 1999; 99EP-0200216.
PR
XX

PA (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX

PI Cornelis S, Beyaert R;
XX
DR WPI; 2000-499331/44.
XX
PT Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -
XX
PS Disclosure; Page 32; 57pp; English.

xx The PYSLRE protein kinase gene can be translated to give two isoforms,
cc
cc
cc p110 and p58. Transcription of p58 mRNA was found to start at an

CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence is the corresponding RNA of the PITSLRE
CC trans.

res described in AAV4/3/10. The IRES sequence and fragments of it may be used to induce cell cycle dependent initiation of translation in eukaryotic cells. Vectors containing the IRES may be used for the propagation of a variety of proteins.

Sequence 732 BP. 77 A, 36 C, 80 G, 20 H, 0 other.
SD
CC
CC
CC
CC
XX

Query Match	33 6A	Score 222	DR 21	Length 222
DR Sequence 222	DF	77 A	50 C	65 G
			20 U	0 Other

Best Local Similarity	91.08	Pred. No. 4	4-40	
Matches 202; Conservative	20	Mismatches	0	Indels 0; Gaps 0;

[illegible][illegible]

Dd	61	ucaggnucugcgaaagaagcagacgagcaggaggaagcagacccacaguu	120
Oy	559	gaagaaatcaagaggaagaaagcagaaagcagagcagaggaagaccgcacqacaactctgaqq	618

Db 121 gaagaaucagagagagagagagagagagagagagagacgcgcagcaacucugagag 180

QY	619	gacaccagcagctcttcgcgaagaatcaatcgaggagaattg	660
Dd	181	gcaccagagccagucugccgaagaaguagggagaagaauy	222

RESULT 6

AAA/3715	standard; DNA; 87 BP.
ID	AAA73715
XX	

AC	AAA/3/15;
XX	
DT	07-DEC-2000 (first entry)
XX	

Functional fragment of human PITSLRE protein kinase gene IRES.

XX
KW
LM
NM
LNBS,
cancer;
XX

Homologous recombination; p58; p110; protein kinase; ds.
gene therapy; human; gene therapy;

CS homo sapiens.
 XX
 PN WO200044896-A1.
 XX

03-AUG-2000.
PD
XX
26-JAN-2000: 2000WO-EP00543.

XX
PR 26-JAN-1999; 99EP-0200216.
XX

PA (VLA-) VLAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
I Cornelis S, Bevaert R;

XX
DR
XX WPI; 2000-499331/44.

Db 709 cgcgaaggagcgggagcccgcaaggagtgctgcacatcacccgaacgatgagagagac 768
Qy 301 tacagcgacaaagtgaagccagccacttgatgcagccgcctcgccgcgcggag 360
Db 769 tacagcgacaaagtgaagccagccacttgatgcagccgcctcgccgcgcggag 828
Qy 361 cgggtcgaattggagacgcccgaagccagtaaaagaagaataatggagaagaagac 420
Db 829 cgggtcgaattggagacgcccgaagccagtaaaagaagaataatggagaagaagac 888
Qy 421 ctgctgtcgcactacagacatcagcgacagcgagaggaagaccagctcgccagatcc 480
Db 889 ctgctgtcgcactacagacatcagcgacagcgagaggaagaccagctcgccagatcc 948
Qy 481 tcgtccacagaatcagctcgaagtgcttgaggaagaagagagagagagagagag 540
Db 949 tcgtccacagaatcagctcgaagtgcttgaggaagaagagagagagagagagag 1008
Qy 541 gaggaagggaagcagcagtgaaagacagagagagagagagagagagagagagac 600
Db 1009 gaggaagggaagcagcagtgaaagacagagagagagagagagagagagagagac 1068
Qy 601 ggcagcaactctgagagagcagcagcagctcgcgaagaagtaagtgaaggaagac 660
Db 1069 ggcagcaactctgagagagcagcagcagctcgcgaagaagtaagtgaaggaagac 1128

RESULT 3
AAA73713
ID AAA73713 standard; DNA; 468 BP.

XX AAA73713;

Dt 07-DEC-2000 (first entry)

XX 5' extension of PITSLRE protein kinase internal ribosome entry site.

DE IRES: internal ribosome entry site; PITSLRE; human; gene therapy;

KW cancer; restenosis; p58; p110; protein kinase; ds.

XX Homo sapiens.

XX WO200044896-A1.

XX 03-AUG-2000.

XX 26-JAN-2000; 2000WO-EP00643.

XX 26-JAN-1999; 99EP-0200216.

XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX Cornelis S, Beyaert R;

XX WPI: 2000-499331/44.

XX Nucleic acids encoding internal ribosome entry sequences useful for

PT directing protein expression in gene therapy procedures -

XX Claim 9; Page 33; 57pp; English.

XX The PITSLRE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence contains a 5' fragment of the PITSLRE
CC protein kinase IRES and nucleotides upstream of the IRES. The IRES
CC sequence and fragments of it may be used to induce cell cycle
CC dependent initiation of translation in eukaryotic cells. Vectors
CC containing the IRES may be used for the preparation of compositions for
CC the treatment of and/or prevention of a disease by gene therapy. Such
XX diseases may be cancer and restenosis.

SQ Sequence 468 BP; 144 A; 100 C; 182 G; 42 T; 0 other;

Query Match 70.9%; Score 468; DB 21; Length 468;

Best Local Similarity 100.0%; Pred. No. 5; Be-94; Mismatches 0; Gaps 0;

Matches 468; Conservative 0; Indels 0;

Qy 1 cccgaagtcggaagacatcgagaagacagagataaagctcgccgggaatggaaaga 60
Db 1 caagaaagcgcgaagacacatcgagaagacagagataaagctcgccgggaatggaaaga 60
Qy 61 cagaagaagaggaatggcaagggagcattccagagagagaaaggggaatgagcg 120
Db 61 cagaagaagaggaatggcaagggagcattccagagagagaaaggggaatgagcg 120
Qy 121 tgcctcttaaggagccgcttgagcaggttagaagaagcgagacggagagcgagatg 180
Db 121 tgcctcttaaggagccgcttgagcaggttagaagaagcgagacggagagcgagatg 180
Qy 181 cgggaagcagaagaagagcagcgagcagaagagcgagcggcgagcgagagagcg 240
Db 181 cgggaagcagaagaagagcagcgagcagaagagcgagcggcgagcgagagagcg 240
Qy 241 cgcgaagagcgagagcgcccgcaagggagtgctgcacatcacccgaacgagtgagagagac 300
Db 241 cgcgaagagcgagagcgcccgcaagggagtgctgcacatcacccgaacgagtgagagagac 300
Qy 301 tacagcgacaaagtgaagccagccacttgagtcgagcccgctcgccgcgcggag 360
Db 301 tacagcgacaaagtgaagccagccacttgagtcgagcccgctcgccgcgcggag 360
Qy 361 cgggtcgaattggagacgcccgaagccagtaaaagaagaataatggagaagaagac 420
Db 361 cgggtcgaattggagacgcccgaagccagtaaaagaagaataatggagaagaagac 420
Qy 421 ctgctgtcgcactacagacatcagcgacagcgagagagagagagac 468
Db 421 ctgctgtcgcactacagacatcagcgacagcgagagagagagagac 468

RESULT 4

AAA73709
ID AAA73709 standard; DNA; 222 BP.

XX AAA73709;

Dt 07-DEC-2000 (first entry)

XX Human PITSLRE protein kinase gene internal ribosome entry site.

XX IRES: internal ribosome entry site; PITSLRE; human; gene therapy;

KW cancer; restenosis; p58; p110; protein kinase; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FT misc_signal 1..222

FT /tag= a

FT /label= IRES

FT /note= "Internal ribosome entry site"

XX WO200044896-A1.

XX 03-AUG-2000.

XX 26-JAN-2000; 2000WO-EP00643.

XX 26-JAN-1999; 99EP-0200216.

XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX Cornelis S, Beyaert R;

XX Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -
XX
XX
PS Claim 10; Page 34; 57pp; English.

Query Match	100.0%;	Score 660;	DB 21;	Length 660;
Best Local Similarity	100.0%;	Pred. No. 4.8e-136;		
Matches 660; Conservative	0;	Mismatches	0;	Gaps 0;

OY	1	caagaaacgtctcgagaaataagatactcgaaataaataaagaataaagatctcgccgagatctgagaaaga	60
Db	1	caacgaataagcttcgagaaatacgaacatctgaaataaataaagaataaatacctcgccgagatctgagaaaga	60
OY	61	caagaaagaaagagaaatactgacaagagagagatctccagtcgagaaagagagagagaaatctgacgtcg	12
Db	61	caagaaagaaagagaaatactgacaagagagagatctccagtcgagaaagagagagagaaatctgacgtcg	12
OY	121	tgacctcttcagggagaccgcgtctcgagagacagtctaaagaagagagacgaggagctggagctgacaagatcg	18
Db	121	tgacctcttcagggagaccgcgtctcgagagacagtctaaagaagagagacgaggagctggagctgacaagatcg	18
OY	181	cggagagacagacagaagagagacgacgcggagacagaagaagagacgcgacgcggcgagctcgagagagcg	24
Db	181	cggagagacagacagaagagagacgacgcggagacagaagaagagacgcgacgcggcgagctcgagagagcg	24
OY	241	cgcagaagagagagagagagaccgcgcagagaaatctgctcacatccagaaatacgaatacgaagaagagac	30
Db	241	cgcagaagagagagagagagaccgcgcagagaaatctgctcacatccagaaatacgaatacgaagaagagac	30
OY	301	tacaacgcagacaagaatctgaaagaacacagaccacactctgagatctgcagaccgcgcctcgccgcgcgcgagag	36
Db	301	tacaacgcagacaagaatctgaaagaacacagaccacactctgagatctgcagaccgcgcctcgccgcgcgcgagag	36
OY	361	cggatctcgagatctggagagacggccggagagaccgaataaagaagagaaatctgagaagaagagac	42
Db	361	cggatctcgagatctggagagacggccggagagaccgaataaagaagagaaatctgagaagaagagac	42
OY	421	ctgctgtgcgcgactttaaagagacatacagccagacagcggagagagacacagctctggccagcttc	48
Db	421	ctgctgtgcgcgactttaaagagacatacagccagacagcggagagagacacagctctggccagcttc	48
OY	481	tcgcagacagaatacagagccacagatctctcagagaaagagagagagagagagagagagagagag	54
Db	481	tcgcagacagaatacagagccacagatctctcagagaaagagagagagagagagagagagagagagag	54
OY	541	gagagaaagagagacacacagctgaagaatacagaaagagagagagagagagagagagagagagac	60
Db	541	gagagaaagagagacacacagctgaagaatacagaaagagagagagagagagagagagagagagac	60
OY	601	ggcagacaaactctgagagagagatcagaaagacagctctgcgaaagaagaatgaagtctgagagaagaatcg	66
Db	601	ggcagacaaactctgagagagagatcagaaagacagctctgcgaaagaagaatgaagtctgagagaagaatcg	66
RESULT 2			
ID	AAA73711		
XX	AAA73711 standard; DNA; 2471 BP.		

AC	AAA73711;
XX	
DT	07-DEC-2000 (first entry)
XX	
DE	Human PITSLRE protein kinase isoform alpha2-2.

PN WO2000044896-A1.
XX
XX
PD 03-AUG-2000.
XX
XX PF 26-JAN-2000; 2000WO-EP00643.
XX
XX 26-JAN-1999; 99EP-0200216.
PR

XX directing protein expression in gene therapy procedures -
 PT Nucleic acids encoding internal ribosome entry sequences useful for
 ZN
 PS Disclosure: Page 32-33, 57pp; English.

The pTRK6 protein kinase gene can be translated to give two isoforms p110 and p56. Transcription of p58 mRNA was found to start at an internal ribosome entry site (IRES). The IRES element was found to be cell cycle regulated, p58 is produced during the G2/M stage of the cell cycle. The present sequence is the human p135Rb protein kinase isoform alpha-2. This sequence contains the IRES. The IRES sequence and fragments of it may be used to induce cell cycle dependent initiation of translation in eukaryotic cells. Vectors containing the IRES may be used for the preparation of compositions for the treatment of and/or prevention of a disease by gene therapy. Such diseases may be cancer and restenosis.

Sequence 2471 BP; 697 A; 601 C; 809 G; 364 T; 0 other;

Query Match	100.0%;	Score 660;	DB 21;	Length 2471;
Best Local Similarity	100.0%;	Pred. No. 6.4e-136;		
Matches 660; Conservative	0;	Mismatches	0;	Gaps 0;

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 19:03:42 ; Search time 355.05 Seconds
(without alignments)
1593.676 Million cell updates/sec

Title: US-09-915-060-6
Perfect score: 660
Sequence: 1 cagcagctcggaacgaca.....aagtaagtgcggaagaatg 660

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: N.Geneseq_1101.*
2: /SIDS2/gcgcdata/geneseq/geneseqn/NA1980.DAT.*
3: /SIDS2/gcgcdata/geneseq/geneseqn/NA1981.DAT.*
4: /SIDS2/gcgcdata/geneseq/geneseqn/NA1982.DAT.*
5: /SIDS2/gcgcdata/geneseq/geneseqn/NA1983.DAT.*
6: /SIDS2/gcgcdata/geneseq/geneseqn/NA1984.DAT.*
7: /SIDS2/gcgcdata/geneseq/geneseqn/NA1985.DAT.*
8: /SIDS2/gcgcdata/geneseq/geneseqn/NA1986.DAT.*
9: /SIDS2/gcgcdata/geneseq/geneseqn/NA1987.DAT.*
10: /SIDS2/gcgcdata/geneseq/geneseqn/NA1988.DAT.*
11: /SIDS2/gcgcdata/geneseq/geneseqn/NA1989.DAT.*
12: /SIDS2/gcgcdata/geneseq/geneseqn/NA1990.DAT.*
13: /SIDS2/gcgcdata/geneseq/geneseqn/NA1991.DAT.*
14: /SIDS2/gcgcdata/geneseq/geneseqn/NA1992.DAT.*
15: /SIDS2/gcgcdata/geneseq/geneseqn/NA1993.DAT.*
16: /SIDS2/gcgcdata/geneseq/geneseqn/NA1994.DAT.*
17: /SIDS2/gcgcdata/geneseq/geneseqn/NA1995.DAT.*
18: /SIDS2/gcgcdata/geneseq/geneseqn/NA1996.DAT.*
19: /SIDS2/gcgcdata/geneseq/geneseqn/NA1997.DAT.*
20: /SIDS2/gcgcdata/geneseq/geneseqn/NA1998.DAT.*
21: /SIDS2/gcgcdata/geneseq/geneseqn/NA1999.DAT.*
22: /SIDS2/gcgcdata/geneseq/geneseqn/NA2000.DAT.*
23: /SIDS2/gcgcdata/geneseq/geneseqn/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	660	100.0	660	21	AAA73714 Human PITSURE prot
2	660	100.0	2471	21	AAA73711 Human PITSURE prot
3	468	70.9	468	21	AAA73713 5' extension of PI
4	222	33.6	222	21	AAA73709 Human PITSURE prot
5	222	33.6	222	21	AAA73710 RNA of human PITSURE
6	87	13.2	87	21	AAA73715 Functional fragmen
7	87	13.2	575	22	AAI23315 Probe #13248 for g
8	87	13.2	575	22	AAI48636 Probe #17322 used t
9	87	13.2	575	22	AAI08956 Probe #8947 used t
10	87	13.2	1969	22	AAI14105 Probe #4038 for ge
11	87	13.2	1969	22	AAI35486 Probe #4172 used t

12	87	13.2	1969	22	AAI03958
13	81.6	12.4	2803	22	AAI07132
14	81.2	12.3	2805	22	AAI07130
15	79.6	12.1	2800	22	AAI07131
16	78.6	11.9	1925	20	AAI90924
17	78.4	11.9	1926	21	AAI50254
18	78.4	11.9	1926	22	AAI82902
19	78.4	11.9	2580	21	AAI74544
20	78.4	11.9	5452	20	AAI90923
21	78.4	11.9	8705	20	AAI23778
22	78.4	11.9	9600	19	AAI21683
23	78.4	11.9	10380	20	AAI22248
24	78.4	11.9	10596	14	AAO51731
25	78.4	11.9	10596	17	AAI40348
26	78.4	11.9	10596	20	AAI15650
27	78.4	11.9	16080	21	AAI55553
28	75.4	11.4	3489	21	AAI30290
29	75.4	11.4	3489	22	AAI82901
30	75.4	11.4	32207	20	AAI73805
31	75.4	11.4	137507	19	AAI19941
32	75.2	11.4	795	19	AAI55830
33	75.2	11.4	799	19	AAI55831
34	66.8	10.1	49999	20	AAI23891
35	66.8	10.1	49999	20	AAI23896
36	65.4	9.9	49999	20	AAI23901
37	65	9.8	15672	12	AAO10613
38	64.6	9.8	2643	19	AAO39212
39	64.6	9.8	3717	21	AAI64660
40	64.2	9.7	470	22	AAI12522
41	64.2	9.7	470	22	AAI13875
42	64.2	9.7	470	22	AAI02430
43	63.6	9.6	726	22	AAI22214
44	63.6	9.6	726	22	AAI47511
45	63.6	9.6	726	22	AAI07913

ALIGNMENTS

RESULT 1
ID AAA73714 standard; DNA; 660 BP.
AC AAA73714;
DT 07-DEC-2000 (first entry)
DE Human PITSURE protein kinase gene IRES and upstream sequence.
XX IRES: internal ribosome entry site; PITSURE: human; gene therapy;
KW cancer; retinosis; p58; p110; protein kinase; ds.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT misc_signal 439..660
FT /*tag= a
FT /label= IRES
FT /note= "Internal ribosome entry site"
XX
XX WO200044896-A1.
XX
XX 03-AUG-2000.
XX
XX 26-JAN-2000; 2000WO-EP00643.
XX
XX 26-JAN-1999; 99EP-0200216.
XX
XX (VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
XX Cornelis S, Beyaert R;
XX
XX WPI: 2000-499331/44.

Probe #3949 used t
Canine retinitis p
Canine retinitis p
Canine retinitis p
Epstein Barr Virus
Epstein Barr Virus
Nucleotide sequenc
Anti-sense strand
Vector pshuttle DN
Vector plasmid pcm
Nucleotide sequenc
Plasmid pcISBON f
Plasmid pcISBON f
Nucleotide sequenc
DNA clone pCK C1.
Kaposi's sarcoma-a
Nucleotide sequenc
KSHV LTR DNA (nucl
KSHV long unique c
FLGA insert stabl
Nucleotide sequenc
Murine LOBO genom
Murine LOBO homolo
Human LOBO homolo
Rianodin receptor
CENP-B CDNA. Homo
DNA encoding cent
Probe #2455 for ge
Probe #2561 used t
Probe #421 used t
Probe #12147 for g
Probe #16197 used t
Probe #7904 used t

Db	576	CAGAAGACAAGGAAATATGCAAGAGGAGCATTTCCAGGAGAGAA-----	61/
Oy	121	tgcccttcaggagccgcttgagcaarttagaaagaaacgagagcggaacgcaaatg	180
Db	618	-----AGGAGCCGCTTTGGACAGCTTTAAAGAGGAGCGGAGCGGACGCAAGATG	668
Oy	181	cggagcagcagaagaaagcaacgagagcagaagaagacgagacgagcgagcgagagcg	240
Db	669	CGGAGCGAGCGAAGAGGAGCAGCGGAGCAGAAAGAGCGCGAGCGGCGGAGAGCGG	728
Oy	241	cgcgaagagcgaggagcccgagaggaaatgctctcacatcaacgaatgaagaagac	300
Db	729	CGCAAGAGCGGAGAGCCCGCAGAGAACTGCTGCATCACCGAACATGAGAGGAC	788
Oy	301	tacacgcacaaagtgaagaccagccactgagtcgcaagcccgctctggcgcgcgagg	360
Db	789	TACACGCACAAGTAGAAGCCAGCCACTGAGTGTGCACCGCCGCTCGGCGGGAG	848
Oy	361	cggctcgagcttgaggagacgagccggaagccagtaaaagaagaagaatgaaagaagac	420
Db	849	CGGTTCCGAGTTGGGAGAGCGCGCGAAGCCAGTAAAGAGAGAAATGAGAAAGGAC	908
Oy	421	ctgcgtccgacattacagacacatcagcagacagcgagaaagaccagctcgcgcgagtc	480
Db	909	CTGCTGTCCGACTTACAGAGACATCAGCCACAGAGAGAAAGACACAGCTCGGCGAGTCC	968
Oy	481	tgcctcagcagaatcagctcaggtctctgaagaaagaagagagagagagagagag	540
Db	969	TGCTGACGACGATCAGCTCAGGTTCTGAGGAAGAGAGAGAGAGAGAGAGAG	1028
Oy	541	gaggaagagagagacacagcaggaagaatca-----gagagagagagagagagagag	591
Db	1029	GAGGAAGGAGGACACAGTGAAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAG	1088
Oy	592	gagagagaccgagcagaactctgaagagagcatcagacagctcgcgcaagaagtgaag	651
Db	1089	GAGGAGACCGGACGACACTGAGAGGACATCAGAGCAGTCTGCCGAGAAGTAGTGAG	1148
Oy	652	gaaagaatg 660	
Db	1149	GAAGAAATG 1157	
RESULT 15			
LOCUS	AF067516		
DEFINITION	Homo sapiens PITSIRE protein kinase alpha SV10 isoform (CDC2L1)		
ACCESSION	AF067516		
VERSION	AF067516		
KEYWORDS	AF067516.1 GI:3850311		
SOURCE	.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
	1 (bases 1 to 2533)		
	Gururaira, R., Lahti, J.M., Grenet, J., Easton, J., Gruber, I.,		
	Ambros, P.F. and Kidd, V.J.		
	Duplication of a genomic region containing the Cdc2L1-2 and		
	MM21-22 genes on human chromosome 1p36.3 and their linkage to D12Z		
	Genome Res. 8 (9), 929-939 (1998)		
	98424414		
	2 (bases 1 to 2533)		
	Gururaira, R., Lahti, J.M., Grenet, J., Easton, J., Gruber, I.,		
	Ambros, P. and Kidd, V.J.		
	Direct Submission		
	Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's		
	Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA		
	Location/Qualifiers		
	1..2533		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="11"		
FEATURES			
	source		

QY	121	tgctcttcacggagccgcttgagacggttagaagaagcggagacggagacggaatg	180
QY	181	cggagacgaacgaaggagacgacggagacgaagaagcggagacggagacggagacggag	240
Db	655	CGGAGCAGCAGAGAGAGCAGCGCGGAGCAGAGGACCGCAGCGGGGGGGGAGGACGG	714
QY	241	cgcgaagagcggagagcccgacggaggaagtgtctcaccatcacgaaacgaatgaagagac	300
Db	715	CGCAGAGACGGGAGGCGCGGAGGGAAGTGCTGCATCATCCGACGATGAGAGAGAGAC	774
QY	301	tacagagacaagaatgaaagccagccacbtgaatgcgagacgcgcgtgcgcgcgcggagag	360
Db	775	TACAGGCAAAAGTGAAGCCAGCCACTGGAGTCCGAGCCCGCTTCGGCCCGCGGAGG	834
QY	361	cggcttcagcttgaggagacggccggaagccagataaagaagagaaaaatgaagaagaagac	420
Db	835	CGGTTGCAATTTGGAGACGCGCGGAAGCCAGTAAAGAGAGAAATGGAGAAAGGAGAC	894
QY	421	cgcgtctcgcgacttaagagacatcagcgacgacgagagaaagacacgctcgcgcgagttcc	480
Db	895	CTGCTTCCGACTTTACAGGACATCAGCGCAGCGAGAGAGAACCCAGCTCGGCCGATGCC	954
QY	481	tcgtcagcagaaatcaggtccaggttcttgaggaaagagagagagagagagagagagag	540
Db	955	TGCTCAGCAGATCAGAGGCTCAGGTTCTGAGGAAGAGAGAGAGAGAGAGAGAGAGAG	1014
QY	541	ggaggagagagagacacagatgaagaatcagaagagagagagagagagagagagagacc	600
Db	1015	GAGGAAGGGAGAGACAGTGAAGAAATCAGAGGAGAGAGAGAGAGAGAGAGAGAGAC	1074
QY	601	ggcagcaactcgaagagacatcagagcagctctgcggaagagatgaatgaagaagaatg	660
Db	1075	GGCAGCAACTCAGAGAGGACATCAGAGCAGTCTCCGAGGAAGTAACTGAGAGAGAAATG	1134
RESULT	13		
LOCUS	AF067512	2486 bp	07-NOV-1998
DEFINITION	Homo sapiens PTPKLE protein kinase alpha SV1 isoform (CDC2L1)		
ACCESSION	AF067512		
VERSION	AF067512.1	GI:3850303	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 2486)		
TITLE	Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P.F. and Kidd,V.J.		
JOURNAL	Duplication of a genomic region containing the Cdc2L1-2 and MWP21-22 genes on human chromosome 1p36.3 and their linkage to D12Z2		
REFERENCE	Genome Res. 8 (9), 929-939 (1998)		
AUTHORS	2 (bases 1 to 2486)		
JOURNAL	Gururajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P. and Kidd,V.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's		
FEATURES	Research Hospital, 332 N. Lauderdale st., Memphis, TN 38105, USA		
SOURCE	Location/Qualifiers		
	1..2486		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/map="1p36.3"		
gene	1..2486		
	/gene="CDC2L1"		
CDS	112..2460		
	/gene="CDC2L1"		
	/note="corresponding genomic sequence deposited as GenBank		

[illegible]

QY	121	tgctcttcaaggagacgcgtcttgagcaqttagaagaagcgaggcgagagcgcaaga	tg	180
QY	571	-----AGGGACCGCGTTGGAGCAGTTAGAAAGGAAGCGGGAGCGGACGCAAGTG		621
QY	181	cgggagcagcagaagaagcagcgaggagcaagaagcgcgagcgcgagcgagagcg		240
Db	622	CGGGAGCACCAACAAGAGAGCGGGAGAGAGAAGAGCGCCGAGCGCGGCGAGACGG		681
QY	241	cgcagaagcgaggagcccgcaagggaagtgtctgaca tcaaccgaagatagagagac		300
Db	682	CGCAAGAGACGGGAGGCCCGCAGGGAATGTCTGCACATTCACCGAGCATGAGAGGAC		741
QY	301	taccagcgacaagtgaaagccagccacatgagctgcagagcccgctcgcgcgcgagag		360
Db	742	TACACGCAAAAGTGAAGCCAGCCACTGAGATCGACGCCCGCTCGGCGCGGGAG		801
QY	361	cggttcgagtgtagagagcgccggagagcccaataaagaagaagaataatgaaagaagagac		420
Db	802	CGGTTCCGATGTGGGAGACGCGCGGAGGCGAGTAAAGAGAAATGCAAGAAAGGAC		861
QY	421	ctgctgtccgaattacagagacatcaagcgacagcgagaaagacacagctcgccgagtc		480
Db	862	CTGCTGTCGCACTTACAGAGCATCGCAGCGAGGAGGAGACACAGCTGGCGGACTCC		921
QY	481	tcttagaagaatcagagctcaggtcttgaagaaagagagagagagagagagagag		540
Db	922	TGCTAGCAGATCAAGCTCAAGTTCTGAGGAAGAGAGAGAGAGAGAGAGAGAGAG		981
QY	541	gaggaagagagacacacagtgaaagaaatcaagagagagagagagagagagagagac		600
Db	982	GAGGAAGGAGACACAGTGAAGAAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAC		1041
QY	601	ggcagcaactctgagagagcatcagagcagctctgcgaagaagttaagtggaaagatg		660
Db	1042	GCGACCACTCTGAGGAGGATCATGAGAGCAGTCTCCGAGAACTAATGAGGAAGAAATG		1101
RESULT	11			
LOCUS	HSU07704	2471 bp	MRNA	PR1
DEFINITION	Human protein kinase PITSLRE isoform PBETA21 (P58CD211) mRNA,			
ACCESSION	U07704			
VERSION	U07704.1	GI:507426		
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
TITLE	1 (bases 1 to 2471)			
JOURNAL	Xiang,J., Lohlt,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.			
MEDLINE	Molecular cloning and expression of alternatively spliced PITSLRE			
REFERENCE	J. Biol. Chem. 269, 15786-15794 (1994)			
AUTHORS	J. Kidd, V.J.			
TITLE	2 (bases 1 to 2471)			
JOURNAL	Kidd, V.J.			
FEATURES	Direct Submission			
SOURCE	Submitted (14-MAR-1994) Vincent J. Kidd, St. Jude Children's			
	Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St.,			
	Memphis, TN 38101, USA			
	Location/Qualifiers			
	1..2471			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="1"			
	/map="1p36-2"			
	/cell_line="Hela S3"			
	/cell_type="epitheloid"			
	/tissue_type="cervix"			
	/clone_lib="Hela CDNA library"			
	/dev_stage="adult"			

Db	434	CACGACGCTGGAAACGACATCGGAAGAACACAGATTAAGCTGCCCGGAAATGGGAAGA	493
QY	61	cagaagaagaaagaaatgycagaaggagcatccagsgagaagaaggggaatgatlgtgcgtg	120
Db	494	CAGAAGAGAAAGGGAATGTGCGAAGGAGCAATTCCTCGAGAGAGA-----	535
QY	121	tgctcttcagggagccgcttggagcagtlgtgaagaagcgggagcgggaacggaagatg	180
Db	536	-----AGGGACCCCTTGAGCAGCTTAAAGAGAAAGCCGGGAGCGGAGCCGCAAGTGG	586
QY	181	cgaggagcagcagaagggagcagcgggagcagaagaagcgcgagcggcgggcggaagacg	240
Db	587	CGGAGACGACGAGAAAGAGACACGGGAGACAGAAAGAGACGCGGAGCGGCGGAGAGCGG	646
QY	241	cgcagaagcgcgggaagcccgcaaggaaagtgtctgcacatcacgcgaacgatlgaagagac	300
Db	647	CGCAAGGAGCGGAGAGCCCGCAGGGAAGTGTCTGCACATCACCGAACATGAGAGAGAC	706
QY	301	tacagcgaacaagtgtgaaagccagccactggaggtgcgagcccgcttcggccgcccggag	360
Db	707	TACAGCGCAAAAGTGAAGAACCCAGCCACTGGAGTGCACACCCGCTCGGCCCCGGGAG	766
QY	361	cggttcaggttggggagacgcgcgggaagccagtaaaagaagaataatgaaagaagagac	420
Db	767	CGGTTCCGAGTTGGGAGACGCGCGGAGACCCAGTAAAGAGAAAGAAATGAGAAAGGAGAC	826
QY	421	ctgctgtccgaacttacagaagacatcaagcgcacagcggaaggaaacccagctcgccgagtc	480
Db	827	CTGCTGTCCGACTTACAGAGCATCAGCGACAGCGAGAGAGAACCCAGCTCGGCCCACTCC	886
QY	481	tcgtcagcagaatacagctcaggtctctgaagaagaagggagggaggaagaagagagag	540
Db	887	TCTGTACGAGAGATCAGGCTCAGGTTCTGAGGAAGAGAGAGAGAGAGAGAGAGAGAG	946
QY	541	gaggaaagggaacacacagcgaagaatcagaagagagaagaaggaaagagagagagacc	600
Db	947	GAGGAGGAGGACACCACTGAGAAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGACC	1006
QY	601	ggcagcaactctbgaaaggcatcagaagcagctctgccgaagaagtgaatgtaggaagaatg	660
Db	1007	GGCACCAACTCTGAGAGAGCATCAGACGACGTCTCCCGAAGAAGTAAGTGAGGAAAGAAATG	1066
RESULT	9		
AF067519	AF067519	2439 bp	07-NOV-1998
LOCUS		mRNA	
DEFINITION	Homo sapiens PITSLR protein kinase beta SV1 isoform (CDC2L2) mRNA,		
ACCESSION	AF067519	complete cds.	
VERSION	AF067519.1	GI:3850317	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 2439)		
JOURNAL	Gururajan,R., Lanthi,J.M., Grenet,J., Easton,J., Gruber,I.,		
MEDLINE	Ambros,P. and Kidd,V.J.		
REFERENCE	2 (bases 1 to 2439)		
AUTHORS	Gururajan,R., Lanthi,J.M., Grenet,J., Easton,J., Gruber,I.,		
TITLE	Ambros,P. and Kidd,V.J.		
JOURNAL	Direct Submission		
FEATURES	Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's		
source	Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA		
	Location/Qualifiers		
	1..2439		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		

gene	1.. 2439
CDS	/gene="CDC2L2" 80..2413 /gene="CDC2L2" /note="Corresponding genomic sequence deposited as GenBank Accession Numbers AF080689-AF080697 and AF092426-AF092428" /codon_start=1 /product="P1RSURE protein kinase beta SVL isoform" /protein_id="AAC72084.1" /db_xref="GI:3850318"
Query Match	90.3%; Score 596; DB 9; Length 2439;
Best Local Similarity	95.9%; Pred. No. 8.3e-101;
Matches 633; Conservative	0; Mismatches 0; Indels 27; Gaps 14;
BASE COUNT	695 a 592 c 790 g 362 t
ORIGIN	
Query	1 cagaacgcttggaagacacatcgaaagaacgaataagctcgcggaatgaggaaagaa 60
Db	464 CACGAACGTCGGAACGACATCGAAGAGACGGTTAAAGCTCCCGGAAATGGAAAGA 523
Query	61 cagaagaagaggaaatggaagaggacatccagaagagagaaggggaatgagtcgtg 120
Db	524 CAGAAGAGGAAGGAATGGCAAGGAGCATTCACAGAGAGA----- 565
Query	121 tgcctctcaggagaccgcttggagcagttagaagaagcgggagcgggagcgcaagatg 180
Db	566 -----AGGACCGCTTGGAGCAGTTAAGAAAGGAGCGGAGCGGAGCCGAACATG 616
Query	181 cgggagacagaagaagaagcagcgggagcagaagggcgcgcgcgcggaagacgg 240
Db	617 CGGAGAGCAGAGAGAGAGACGACGGAGACGAGAGAGCCGAGCCGCGAGAGACGG 676
Query	241 cgcagaagcgaggagagccgcgaggaagtgctgcacatccagacagatgagagagac 300
Db	677 CGCAAGAGCGGGAGCGCCGACAGGAAGTGCTGCAACATCACCGACGATAGAGAGAC 736
Query	301 taacagacaaagtgaagccagccacgtgagtcgcagccgcgcctcgccgcgcgaggag 360
Db	737 TACAGCGACAAAGTGAAGAGCCACCACTGGAGTCCACCCGCTCGGCGCCCGGAG 796
Query	361 cggttcgaatttggaagacggcggaagcagctaaagaagagaaatgaaagaagagac 420
Db	797 CGGTTCCAGTTGGAGACGGCCGGAAGCCAGTAAAGAAAGAAATGAGAAAGAGGAC 856
Query	421 ctgctgtccgacttaacagacatcagcgagcgagcgagaggaagaccagctcgccgagtc 480
Db	857 CTGCTGTCCGACTTACAGGACATCACCGACGAGAGAGAAACACGCTCGGCCGAACTCC 916
Query	481 tcgtcagcagaatcaggctcaggttcttgaggagaagagggagggaggaagagagag 540
Db	917 TCGTCAGAGATCAGGCTCAGGTTCTTGAGGAAGAGAGAGAGAGAGAGAGAGAGAG 976
Query	541 gaggaagagacaccagatgaagatacagagggagaggaaggaagagagagagagac 600
Db	977 GAGGAGGAGACCACTGAAGATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAC 1036

TITLE Ambros, P. and Kidd, V.J.
JOURNAL Direct Submission
FEATURES Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
source location/Qualifiers
1..2323

gene	1..2323
	/gene="CDC2L2"
CDS	1104..2297

OY	1	caacgaacttcggaacacgcacatctcgaaagaacaaggttaaagcttcgcgcgggaatcggaaaga	60
Db	466	CACGAACGTCGGAAACGACATCTGAAGAAACACGGATTAAGCTCCGCGGAATTCGGAAGA	523
OY	61	cagaagagaaggaatctgccaaggagcatctcaagagagaagggtgatgtgctg	120
Db	524	CAGAAGAGAGGGAATTCGCAAGGAGCATTCACGAGACAGAA-----	565
OY	121	tgctcttcagggaacgcgtctggagcagttagaagaagcggagcggagcgcgaatg	180
Db	566	-----AGGAGCCGCTTGAGACAGCTTAAGAAAGAGCGGAGCGGACCGCAAGTG	616
OY	181	cgggagcagcagaaggaacagcggagacagaagaagcggagcggcggaggaagcgg	240
Db	617	CGGGAGCGCGAAGAGACACCGGAGCAGAGAGGAGCCGCGCGCGGAGAGCGG	676
OY	241	cgcgaagcgaggagcccgcaaggaaagtgtctcacataccgaataagatagaagagac	300
Db	677	CGCAAGAGCGCGAGCGCCCGCAGGGAATGTCTGCACATCACCGAATGATGAGAGAGAC	736
OY	301	tacagcgacaagaatgaaagccagccaacttggatctgcacacccgcgctcggccgcggag	360
Db	737	TACAGCGACAAAGTGAAGCCAGCCACTGTGAGTGCACCCGCGCTCGGCCGCCCGGAG	796
OY	361	cggtttcgaattgggagagagcgccggaaagccagtaaaagaagaatatgaaagaagagac	420
Db	797	CGGTTTCGAGTTGGGAGAGCGCGGAGCGCCACTTAAGGAAGAAATGAGAGAAAGGAC	856
OY	421	ctgctgtccggaactacagagacatcagcgaagcgagaggaagagaccaggttcggccgagttcc	480
Db	857	CTGCTGTCCGACTTACAGGACATCAGCGACACGCGAGGAGGAGGACCACTCCGCCGAGATCC	916
OY	481	tcgtcagcagaatcacagctcaggtctctgaagaaagaagagagagagagagagag	540
Db	917	TGCTCAGCAGAAATCAGGCGTCAAGTTTCGAGGAAAGAGAGAGAGAGAGAGAGAG	976
OY	541	gaggaagagagcaccagttgaagaatcagagagagaagaagaagagagagagagacc	600

[illegible]

RESULT	8
AF067521	
LOCUS	AF067521 2409 bp mRNA PRI 07-NOV-1998
DEFINITION	Homo sapiens PITSILE protein kinase beta SV3 isoform (CDCL2L) mRNA.
ACCESSION	AF067521
VERSION	AF067521.1 GI:3850321
KEYWORDS	.
SOURCE	human.

BASE COUNT	686 a	586 c	780 g	357 t
ORIGIN				
Query Match	90.3%;	Score 596;	DB 9;	Length 2409;
Best Local Similarity	95.9%;	Pred. No. 8.3e-101;		
Matches 633;	Conservative	0;	Mismatches	0;
			Indels	27;
			Gaps	1;
Oy	1	caggaagtcggaacgaatcgagagaagacagataaagctgcgcgggaatggaaga	60	

OY	421	ctctgcttcgcagcttacagagcatcaagagcagcgagagggaaagaccactgcgcgcgagcttc	480
Db	671	CTGCTGCTCCGACTTACAGAGCATCAGCGACACGCAAGAGGAAGACCACTCGGCCGACTTC	730
OY	481	tcgtcagagaagatcaaggtcaggtcttcctgagaagaagaagagagagagagagagag	540
Db	731	TCGTCAGCAGAGATCAGCGTCTCAGGTTCTTGAGGAAGAAGAGAGAGAGAGAGAGAGAG	790
OY	541	gaggaagagagacaccatgaaatcagaagagaggaagaagaagagagagagagagacc	600
Db	791	GAGGAAGGAGACACACAGTGAAGAAATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAC	850
OY	601	ggcaagcaactctggaggaagcatcagagcagctgcgcgagaagaatgaatgaaagagag	660
Db	851	GCGACCAACTCTGAGAGAGGACATCAGACGAGCTCTGCCAGAGAGATGAGAGAGAGAAATG	910
RESULT	6		
AF067520			
LOCUS	AF067520	2436 bp	mRNA
DEFINITION	Homo sapiens	PR1SLK protein kinase beta SV2 isoform (CDC2L2)	mRNA,
ACCESSION	AF067520		
VERSION	AF067520.1	GI:3850319	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 2436)		
TITLE	Gururajan, R., Lahti, J. M., Grenet, J., Easton, J., Gruber, I., Ambros, P. F. and Kidd, V. J.		
JOURNAL	Duplication of a genomic region containing the Cdc2L1-2 and MME21-22 genes on human chromosome 1p36.3 and their linkage to D1Z2		
MEDLINE	Genome Res. 8 (9), 929-939 (1998)		
REFERENCE	98424414		
AUTHORS	2 (bases 1 to 2436)		
TITLE	Gururajan, R., Lahti, J. M., Grenet, J., Easton, J., Gruber, I., Ambros, P. F. and Kidd, V. J.		
JOURNAL	Direct Submission		
FEATURES	Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA		
source	Location/Qualifiers		
gene	1. 2436		
CDS	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/map="1p36.3"		
	1. 2436		
	/gene="CDC2L2"		
	/gene="CDC2L2"		
	/note="corresponding genomic sequence deposited as GenBank accession Numbers AF080689-AF080697 and AF092426-AF092428"		
	/codon_start=1		
	/product="PR1SLK protein kinase beta SV2 isoform"		
	/protein_id="AAC72085.1"		
	/db_xref="GI:3850320"		
	/translation="MGDEKDSWKYKTUDELLOEKRRKREDEKAELRLKNSDSDRSK		
	RDSLEGEELRHCMETITRNSPYRREDSMEDRGEDSLAKPPQMSWKRYNHRD		
	BRKKKKHAKVKEKEHRRRRKREEDOKARERWOKRREMAHSHRERKNGDGL		
	FDRLLELERRRERKREDOOKREDEKEREERREKREARERSAHPTMRDED		
	YSDKAKASHMSRPPRRPRRRLGQDQKPYKKEKMERDILSDLOISERKTSASA		
	ESSSSGSGSDEEEDEEEDEEBEBCSTSESEDEEEDEEBEETGSNSDEASQSAEY		
	EEEMSEDERENENLIVPESRDRDSDGSESEAEVEEGETPOSSALTBEDYVPD		
	ALPILKQELPKYLPLAOGQSVEEFQCLNRIEGTYGVYVYAKDKRTDEIVALKR		
	KMEKEEGEPTLSLEINTLILKAQHPNIVYREIVGSGNDKIIVANNVYHDSLSM		
	ETMKQPLPGEVKTLMLOLNGVKLHDNMTLHRDLKSNLLSHAGILKYGGLAR		
	EYSGILKATPVYVYQWYRABELLIGAKEYSTADVMGVCITFELLTORKEPENSE		
	IDOLINKVRELTQSEKTPMPCSELPVKKMKTFSPHYNLRKRRGALLNSQGDLMN		
	KFLVTPGGRISAEDEGLKNEYFRRETPPLDIPDSMPTPAKSEQDQVYKRGTSPPR		
	EGYVSGQDDDLKEGTFHLTTNQGASAGPGFSLE"		

[illegible]

LKMEKEGEPITSLREINTILKAQHNPITVREIVVGSNMDKITVIMNVVHDLKSL
METMKQPLFGEVTKLMTDILRGVKKHLHNMWILHRLDKTNNLLSHAGILKQDFELA
REYSGPLKATPVVVTLMYRAPELLGAKESYVAVMSVGLTCEGGLLOKLPFKS
EIDQINKVEFDLGTPESEKIMPYSELPAVKMFSHPYNNLRKREGALLSPOGPDLM
NKLTFEPGRISAEGLKHEYRETPLEIDSMFPTPAKESQOKVKNKGISPRPEG
GLQISOLGDDDLKETETFHLTTTNOGASAGPFSLEF"

BASE COUNT 697 a 601 c 809 g 364 t

ORIGIN

Query Match 100.0%; Score 660; DB 9; Length 2471;
Best Local Similarity 100.0%; Pred. No. 1.3e-112;
Matches 660; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caggaagcgtcgaaagacacatcgagaagacaagataaagctcgccggagatggagaaga 60
|||||
DB 469 CACGAAGCTCGAAGACATCGAAGAAAGAGATTAAGCTCGCCGGGATGGAAAGA 528
|||||
QY 61 cagaagagaaaggaatgagcaagggagcaltccagagagaagaagggaatgagtcgtg 120
|||||
DB 529 CAGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 588
|||||
QY 121 tgcctcttcagagagccgcttgagagcagtttagaagagaagcggagagcggagagatg 180
|||||
DB 589 TGCCTCTTACGGGAGCCCTTGAGACATTTAGAAAGAGAGCGGAGCGGAGCGCAAGATG 648
|||||
QY 181 cggagagcagagaagagcagcgagagcagaagagcgcgagcgcgagagagcg 240
|||||
DB 649 CGGAGCGACGAGAGGAGGAGCGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 708
|||||
QY 241 cggagaagagcggagagcccgaggaagagtgtctgcacatccagaaacgatgagaagagc 300
|||||
DB 709 CGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 768
|||||
QY 301 taacagcgacaagatgaaagccagccacgtgagtcgagccgctcgccgagcgagag 360
|||||
DB 769 TACAGCGACAAAGTGAAGAGCCAGCCATCGAGTCGAGCCCGCTGGCCGCGGAGAG 828
|||||
QY 361 cggattcgagttggaagacggcggaagccaagtaaaagagaagaatgagaagagagc 420
|||||
DB 829 CGGTTGAGTGGGAGACGGCCGAGAGCCAGTAAAGAGAGAAATGGAAGAAAGGAGAC 888
|||||
QY 421 ctgctcttcagactacagagacatcagcgagcagagagaagaagaccgctcgccgagtc 480
|||||
DB 889 CTGCTGTCCGACTTACAGACATCACGACGAGGAGGAGAGAGCCGCTGCGCGGAGTCC 948
|||||
QY 481 tgcgcagcagaatcaggtctcaggttctgagagaagaagagaagagaagagaagag 540
|||||
DB 949 TCGTTCAGCAAGATCAGGCTCAGGTTCTGAGGAAGAAAGAGAGAGAGAGAGAGAG 1008
|||||
QY 541 gaggagagagagcagcagtaagaatcagagagagagagaagagaagagaagagagacc 600
|||||
DB 1009 GAGGAAGGAGAGCAGCAGTGAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAC 1068
|||||
QY 601 ggcagcaactctgagagagcagcagcagctctgcgaagaagtaagtgaagaagaatg 660
|||||
DB 1069 GGCAGCAACTCTGAGAGAGCAGTACAGAGCAGTCTGCCGAAGAAAGTAAAGTACGAGAAAGATG 1128
|||||

RESULT 5
AF067525 2329 bp mRNA P1 07-NOV-1998
LOCUS Homo sapiens P1TSLRE protein kinase beta SV13 isoform (CDC2L2)
DEFINITION
AF067525 mRNA, complete cds.
ACCESSION
VERSION AF067525.1 GI:3850329
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2329)
AUTHORS Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,

TITLE Ambros, P.F. and Kidd, V.J.
JOURNAL Duplication of a genomic region containing the Cdc2L1-2 and
MEDLINE MRP21-22 genes on human chromosome 1p36.3 and their linkage to D122
REFERENCE Genome Res. 8 (9), 929-939 (1998)
AUTHORS 2 (bases 1 to 2329)
Gururajan, R., Lahli, J.M., Grenet, J., Easton, J., Gruber, I.,
Ambros, P. and Kidd, V.J.
TITLE Direct Submission
JOURNAL Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's
Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
RESEARCH Location/Qualifiers
FEATURES
source
1..2329
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
1..2329
/gene="CDC2L2"
68..1654
/note="Corresponding genomic sequence deposited as GenBank
Accession Numbers AF080689-AF080697 and AF092426-AF092428"
/codon_start=1
/product="P1TSLRE protein kinase beta SV13 isoform"
/protein_id="AAC72090.1"
/db_xref="GI:3850330"

BASE COUNT 634 a 605 c 738 g 352 t

ORIGIN

Query Match 99.5%; Score 656.8; DB 9; Length 2329;
Best Local Similarity 99.7%; Pred. No. 4.9e-112;
Matches 658; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 caggaagcgtcgaaagacacatcgagaagacaagataaagctcgccggagatggagaaga 60
|||||
DB 251 CACGAAGCTCGAAGACGATCGAAGAAAGAGATTAAGCTCGCCGGGATGGAAAGA 310
|||||
QY 61 cagaagagaaaggaatgagcaagggagcattccagagagaagaagggaatgagtcgtg 120
|||||
DB 311 CAGAAGAGAGGAGAAATGCGCAAGGAGAGCATCCAGAGAGAGAAAGGGAGATGATGCTTT 370
|||||
QY 121 tgcctcttcagagagccgcttgagagcagtttagaagagaagcggagagcggagagatg 180
|||||
DB 371 TGCCTCTTCAAGGACCGCTTGAGACAGTTAGAAAGAGAGCGGAGCGGAGCCCAAGATG 430
|||||
QY 181 cggagagcagcaggaagagcagcgagagcagaagagcgagcgcgcgaggaagcag 240
|||||
DB 431 CGGAGCGACGAGAGAGAGCAGCAGGAGAGCAGAGAGAGCGGAGCGGAGAGAGAGCGG 490
|||||
QY 241 cggagaagagcggagagcccgagaggaagtgtctgcacatccagaaacgatgagagagac 300
|||||
DB 491 CGCAAGGAGCGGAGGCGCCGAGGAGGAGTGTGACATCACGGAACGATGAGAGAGAGAC 550
|||||
QY 301 taacagcacaagtgaaagccaagcagctgagtgagtcgagcagccgctcgcgcgcgagag 360
|||||
DB 551 TACAGCGACAAAGTGAAGAGCAGCAGTGAAGTGCAGCCGCCCTTGAGCCGCGGAGAG 610
|||||
QY 361 cggttcgagttggaagacggcgagagccagtaaaagagaagaatgagaagaagagagc 420
|||||
DB 611 CGGTTGAGTGGGAGAGCGCGGAGAGCCAGTAAAGAGAGAAATGGAAGAAAGGAGAC 670
|||||

```
QY 361 cggctcgaattggagacgagccggaagccagtaaaagaagaataatggaaagagagac 420
      |||
Db 823 CGGTCGAGATTGGAGACGCGCGAAGCCAGTAAAGAGAGAAATGCAAGAAAGGAGAC 882
QY 421 ctgctgtccgcacttacagacatcagcagacagcagagagagagacccagctcggccgagttcc 480
      |||
Db 883 CTGCTGTCCGACTTACAGGACATCAGCAGACGCGAGAGAGAACCCAGCTCGGCCGAGTCC 942
QY 481 tgcctcagcagaatcaggtctcgaattcttgaggaagaagaagagagagagagagagag 540
      |||
Db 943 TCGTCAGCAGAAATCAGGCTCAGTCTGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1002
QY 541 gaggaagggagacccagctgaagaatcagagggagagagagagagagagagagagagac 600
      |||
Db 1003 GAGGAAGGGAGCACCGAGTGAAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAC 1062
QY 601 ggagcaaacctcagagagagatcagagcagctgcgcgaagaagtaagttagaggaagag 660
      |||
Db 1063 GGCAGCAACTCTGAGAGGAGCATCAGAGCACTCTGCCGAAGAACTAGTGAAGAGAAATG 1122

RESULT 3
AX033422 2471 bp DNA PAT 21-SEP-2000
LOCUS Sequence 3 from Patent WO0044896.
ACCESSION AX033422
VERSION AX033422.1 GI:10280183
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS 1 (bases 1 to 2471)
TITLE Beyaert, R. and Cornelis, S.
JOURNAL Internal ribosome entry site (IRES), vector containing same and
PATENT: WO 0044896-A 3 03-AUG-2000;
FEATURES
LOCATION/Qualifiers
source 1..2471
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 1..2471
/note="the IRES-activity containing sequence
PITSLRE protein kinase (p110PitSLre) (isoform alpha 2-2)"
BASE COUNT 697 a 601 c 809 g 364 t
ORIGIN
```

```
Query Match 100.0%; Score 660; DB 6; Length 2471;
Best Local Similarity 100.0%; Pred. No. 1,3e-112; Indels 0; Gaps 0;
Matches 660; Conservative 0; Mismatches 0;
```

```
QY 1 cacaagcgtcggaaacgacatcgaagaacagataaagctcgcgggaatgggaaga 60
      |||
Db 469 CACGACGCTCGGAACACGATCGAAGAAACAGATTAAGTCCCGGAGATGGGAAGA 528
QY 61 cagaagaagaagaatggaagagagcatccagagagagaagaaggggaatgagtcgtg 120
      |||
Db 529 CAGAAGAAGAGGAATGGCAAGGAGCATTCAGAGAGAGAAAGGGGGAATGATGCGGTG 588
QY 121 tgcctctcagagacgcttggaagcagttgaaaggaagcggggcgaggaggaagatg 180
      |||
Db 589 TGCCTCTTCAGGACCGCTTGAGACAGTTAGAAAGGAAGCGGAGCGGAGCGCAAGATG 648
QY 181 cgggaagcagcagaagaacagcgggaagcagaagagcgagcgcgagcgagagagagcg 240
      |||
Db 649 CGGGAGCAGCAAGAGAGACACGGGAGCAGAGAGCCGCGGCGCGAGAGAGCGG 708
QY 241 cgcgaagagcgagggcccgagggagagtgctctgcacatccacgaacgagtgaagagac 300
      |||
Db 709 CGCAAGAGCGGAGGAGCCCGCAGGAGAGTGTCTGACATCACGACGATGAGAGAGAC 768
```

```
QY 301 tacacgaacaagtgtgaaagccagccagctgaggtcgacagcccgctctggccgagcgagag 360
      |||
Db 769 TACACGCAACAAGTGTAAAGCCAGCACACTGTGCAAGCCGCTCGGCGCGCGGAG 828
QY 361 cggctcgaattggagacgagccggaagccagtaaaagaagaataatgtgaagaagagac 420
      |||
Db 829 CGGTCGAGATTGGAGACGCGCGAAGCCAGTAAAGAGAGAAATGCAAGAAAGGAGAC 888
QY 421 ctgctgtccgcacttacagacatcagcagcagcagagagagagacccagctcggccgagttcc 480
      |||
Db 889 CTGCTGTCCGACTTACAGGACATCAGCAGACGCGAGAGAGAGAGAGAGAGAGAGAGAG 948
QY 481 tgcctcagcagaatcaggtctcgaattcttgaggaagaagaagagagagagagagagag 540
      |||
Db 949 TCGTCAGCAGAAATCAGGCTCAGTCTGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1008
QY 541 gaggaagggagacccagctgaagaatcagaagagagagagagagagagagagagagac 600
      |||
Db 1009 GAGGAAGGGAGCACCGAGTGAAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAC 1068

RESULT 4
HSU04816 2471 bp mRNA PRI 08-JUL-1994
LOCUS Human protein kinase PITSLRE alpha 2-2 mRNA, complete cds.
DEFINITION U04816
ACCESSION U04816
VERSION U04816.1 GI:507159
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
AUTHORS 1 (bases 1 to 2471)
TITLE Xiang, J., Lahli, J.M., Grenet, J.A., Easton, J.B. and Kidd, V.J.
JOURNAL Molecular cloning and expression of alternatively spliced PITSLRE
MEDLINE 94253170
REFERENCE 2 (bases 1 to 2471)
AUTHORS Kidd, V.J.
TITLE Direct Submission
JOURNAL Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's
Research Hospital, "Tumor Cell Biology, 332 N. Lauderdale St.,
Memphis, TN 38101, USA
FEATURES
LOCATION/Qualifiers
source 1..2471
/organism="Homo sapiens"
/db_xref="taxon:9606"
map_1p36-2"
/clone_lib="Hela S3"
/cell_line="epitheloid"
/tissue_type="cervix"
/clone_lib="Hela CDNA library"
/dev_stage="adult"
1..2471
112..2445
/product="PITSLRE alpha 2-2"
/protein_id="AAA19582.1"
/db_xref="GI:507160"
/translation="MGDEKSKKVVTTDEIIQEKRRKEDEKAKIKRKMSDDRSK
RSLLEGELRDHCMEITITNSPYRRDSKEDRGEEDSLAIKPPQMSRKERYVHHRKD
EKREKKARVKKRERHRRRERHEDOKARERWOKRREMARSHRRRGNDGVC
LPDRLEOLERRERERKREKQOREQERERERERERERERERERERERERERERERERERER
DYSDVKASHMRSPRPPEERPELGGDKPKYKEMERDLDSLQDISDSERTTS
ASSSSAEGSGSEEEEEEEREEESSTSESESESESESESESESESESESESESESESESESE
SEESSEDEERENENHLVPEPSRDRDSGBSEKAEPEVGCCTPOSSALTRGDVPPDS
PALSPLEIKQELPKTLPALQGRSVEEFOCLNRIEETGYVYVYRAKDKTDEIYALKR
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:54:37 ; Search time 2399.05 Seconds
(without alignments)
4538.521 Million cell updates/sec

Title: US-09-915-060-6
Perfect score: 660
Sequence: 1 cagcaacgtcggaaacgaca.....aagtaagtgaagaagaatg 660

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues
Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: gb.ba:*
- 2: gb.htg:*
- 3: gb.in:*
- 4: gb.om:*
- 5: gb.ov:*
- 6: gb.pat:*
- 7: gb.ph:*
- 8: gb.pl:*
- 9: gb.pr:*
- 10: gb.ro:*
- 11: gb.sts:*
- 12: gb.sy:*
- 13: gb.un:*
- 14: gb.vi:*
- 15: em.ba:*
- 16: em.fun:*
- 17: em.hum:*
- 18: em.in:*
- 19: em.om:*
- 20: em.or:*
- 21: em.ov:*
- 22: em.pac:*
- 23: em.ph:*
- 24: em.pl:*
- 25: em.ro:*
- 26: em.sts:*
- 27: em.sy:*
- 28: em.un:*
- 29: em.vi:*
- 30: em.htgo.hum:*
- 31: em.htgo.in:*
- 32: em.htgo.rod:*
- 33: em.htg.hum:*
- 34: em.htg.in:*
- 35: em.htg.rod:*
- 36: em.htg.other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	660	100.0	660	6	AX033425	AX033425 Sequence
2	660	100.0	2465	9	HSU07705	U07705 Human prote
3	660	100.0	2471	6	AX033422	AX033422 Sequence
4	660	100.0	2471	9	HSU04816	U04816 Human prote
5	656.8	99.5	2329	9	AF067525	AF067525 Homo sapi
6	656.8	99.5	2436	9	AF067520	AF067520 Homo sapi
7	596	90.3	2323	9	AF067518	AF067518 Homo sapi
8	596	90.3	2409	9	AF067521	AF067521 Homo sapi
9	596	90.3	2444	9	HSU04817	U04817 Human prote
10	596	90.3	2471	9	HSU07704	U07704 Human prote
11	596	90.3	2477	9	HSU04824	U04824 Human prote
12	596	90.3	2477	9	HSU04824	U04824 Human prote
13	577	87.4	2486	9	AF067512	AF067512 Homo sapi
14	577	87.4	2500	9	AF067514	AF067514 Homo sapi
15	577	87.4	2533	9	AF067516	AF067516 Homo sapi
16	547	82.9	2362	9	AF067523	AF067523 Homo sapi
17	547	82.9	2448	9	AF067522	AF067522 Homo sapi
18	528	80.0	2525	9	AF067515	AF067515 Homo sapi
19	520	78.8	2544	9	AK000081	AK000081 Homo sapi
20	468	70.9	468	6	AX033424	AX033424 Sequence
21	397.8	60.3	3161	10	MUSCDPK	L37092 Mus musculu
22	395.4	59.9	2340	9	HSU04818	U04818 Human prote
23	376.4	57.0	2349	9	AF067517	AF067517 Homo sapi
24	371.4	56.3	1715	9	AF174497	AF174497 Homo sapi
25	252.8	38.3	1936	10	MUSP58GTA	M58633 Mouse p58/G
26	252.8	38.3	2226	9	AF067513	AF067513 Homo sapi
27	222	33.6	222	6	AX033420	AX033420 Sequence
28	222	33.6	222	6	AX033421	AX033421 Sequence
29	166	25.2	110608	9	HS283E3	AL031282 Human DNA
30	165	25.0	297	9	CDCC2L1S06	AF092429 Homo sapi
31	161.8	24.5	297	9	CDCC2L2S06	AF092427 Homo sapi
32	155.4	23.5	592	9	CDCC2L2S08	AF080689 Homo sapi
33	138.4	21.0	712	9	AF067524	AF067524 Homo sapi
34	134.4	20.4	480	9	CDCC2L1S08	AF080678 Homo sapi
35	127.4	19.3	532	9	CDCC2L2S07	AF092428 Homo sapi
36	127.4	19.3	533	9	CDCC2L1S07	AF092430 Homo sapi
37	105.4	16.0	539	9	CDCC2L1S05	AF080688 Homo sapi
38	105.4	16.0	761	9	CDCC2L2S05	AF080697 Homo sapi
39	100	15.2	551	9	AF067527	AF067527 Homo sapi
40	100	15.2	700	9	AF067529	AF067529 Homo sapi
41	99.4	15.1	7218	6	166494	I66494 Sequence 14
42	99.2	15.0	318930	2	AC073495	AC073495 Mus muscu
43	98	14.8	579	9	AF067526	AF067526 Homo sapi
44	98	14.8	728	9	AF067528	AF067528 Homo sapi
45	96.4	14.6	179149	9	AC009785	AC009785 Homo sapi

ALIGNMENTS

RESULT 1
AX033425 AX033425 660 bp DNA PAT 21-SEP-2000
LOCUS Sequence 6 from Patent WO0044896.
DEFINITION AX033425
ACCESSION AX033425
VERSION AX033425.1 GI:10280186
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Beyaert,R. and Cornelis,S.
TITLE Internal ribosome entry site (IRES), vector containing same and
uses thereof
JOURNAL Patent: WO 0044896-A 6 03-AUG-2000;
VLAAMS INTERUNIV INST BIOTECH (BE) ; BEYAERT RUDI (BE) ; CORNELIS
SIGRID (BE)
FEATURES Location/Qualifiers
source 1..660


```

RESULT 14
LOCUS   CNS07C7R 1117 bp DNA GSS 08-JUL-2001
DEFINITION T3 end of clone XBC0AA002F11 of library XBC0AA from strain CBS 767
of Debaryomyces hansenii, genomic survey sequence.
ACCESSION AL438829
VERSION   AL438829
KEYWORDS  AL438829.1 GI:12222242
SOURCE    Debaryomyces hansenii.
ORGANISM  Debaryomyces hansenii.
REFERENCE Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
AUTHORS   Lepingle,A., Casaregola,S., Neuvéglise,C., Bon,E., Nguyen,H.,
          Artiguenave,F., Wincker,P. and Galliardin,C.
          Genomic exploration of the hemiascomycetous yeasts: 14.
          Debaryomyces hansenii var. hansenii
          FEMS Lett. 487 (1), 82-86 (2000)
          20584724
TITLE     2 (bases 1 to 1117)
          Soucier,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
          Boillot-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
          de-Montigny,D., Dujon,B., Durans,P., Lepingle,A., Liorente,B.,
          Malpertuy,A., Neuvéglise,C., Ozler-Kalogeropoulos,O., Potier,S.,
          Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
          Wincker,P. and Weissenbach,J.
          Genomic exploration of the hemiascomycetous yeasts: 1. A set of
          yeast species for molecular evolution studies
          FEMS Lett. 487 (1), 3-12 (2000)
          20584711
JOURNAL   3 (bases 1 to 1117)
MEDLINE   Genoscope.
AUTHORS   Direct Submission
          Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
          2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail :
          seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT   This GSS is part of a random genomic sequencing program of thirteen
          yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
          exiguus, Saccharomyces servazii, Zygosaccharomyces rouxi,
          Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
          lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
          ananatis, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
          Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
          5 kb were prepared and both extremities were sequenced. See
          keywords for description of this sequence and for the sequence of
          the other extremity of this insert.
FEATURES
  source
    1. 1117
      /organism="Debaryomyces hansenii"
      /strain="CBS 767"
      /variety="hansenii"
      /db_xref="taxon:4959"
      /clone="XBC0AA002F11"
      /clone_1lb="XBC0AA"
      /note="end : T3"
      /note=">580"
      /note="similar to Saccharomyces cerevisiae ORF YDR084c [
      similarity to hypothetical C.elegans protein ]"
      /evidence="not_experimental"
BASE COUNT 351 a 159 c 195 g 409 t 3 others
ORIGIN

```

```

Query Match          54.2%; Score 20.6; DB 13; Length 1117;
Best Local Similarity 85.2%; Pred. No. 3.3e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 7 agaggaattcgaagtgaacatttc 33
    1111 1111 11111111 11111
DB 135 AGAGCAATTGATGTCAGCATCTTTT 109

```

```

RESULT 15
LOCUS   AQ402410 546 bp DNA GSS 13-MAR-1999
DEFINITION HS_5060_A1_F10_T7A Rpci-11 Human Male BAC Library Homo sapiens
genomic clone Plate-636 Col-19 Row-K, DNA sequence.
ACCESSION AQ402410
VERSION   AQ402410
KEYWORDS  AQ402410.1 GI:4413322
SOURCE    human.
ORGANISM  Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 546)
          Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
          Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
          Hood,L.
          Sequence-tagged connectors: A sequence approach to mapping and
          scanning the human genome
          Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
          99380589
JOURNAL   High Throughput Sequencing Center
          University of Washington
          401 Queen Anne Avenue North, Seattle, WA 98109, USA
          Tel: (206) 616-3618
          Fax: (206) 616-3887
          Email: jwallace@u.washington.edu
          Clones are derived from the human BAC library Rpci-11. For BAC
          library availability, please contact Plier de Jong
          (plier@dejong.med.buffalo.edu). Clones may be purchased from
          BACPAC Resources (http://bacpac.med.buffalo.edu/ordering-bac.htm)
          or from Resear h Genetics (info@resgen.com). BAC end Web Server:
          http://www.htsc.washington.edu
          Plate: 636 row: K column: 19
          Seq primer: T7
          Class: BAC ends
          High quality sequence stop: 546.
FEATURES
  source
    1. 546
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone="Plate=636 Col=19 Row=K"
      /clone_1lb="Rpci-11 Human Male BAC Library"
      /sex="male"
      /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
      Male blood DNA was isolated from one randomly chosen donor
      and partially digested with a combination of EcoRI and
      EcoRI Methylase. Size selected DNA was cloned into the
      pBACe3.6 vector at EcoRI sites"
BASE COUNT 191 a 95 c 91 g 165 t 4 others
ORIGIN

```

```

Query Match          53.7%; Score 20.4; DB 13; Length 546;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

OY 3 cttcaggaattcgaagtgaacatttc 32
    11111111111111 1111 11
DB 45 CTTCAAGAGCAATTCGACATGTCATCAATT 16

```

```

Search completed: December 4, 2001, 18:14:29
Job time: 6770 sec

```

Location/Qualifiers
1. .541

QY	4	tctcagaggaattcgaagtcacgatacttctg	38
Db	277	ttccggtggagtcgaagtcgccctactgctggccg	243

REFERENCE	Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
AUTHORS	1 (bases 1 to 990) Malpertuy,A., Llorente,B., Blandin,G., Artiguenave,F., Wincker,P. and Dujon,B.
TITLE	Genomic exploration of the hemiascomycetous yeasts: 10.
JOURNAL	Kluyveromyces thermotolerans
MEDLINE	FEMS Lett. 487 (1), 61-65 (2000)
REFERENCE	20584720
AUTHORS	2 (bases 1 to 990) Souclet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Botolin-Pukuhara,M., Bon,E., Brotlier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuvéglise,C., Ozler-Kalogeropoulos,O., Potier,S., Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.
TITLE	Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies
JOURNAL	FEMS Lett. 487 (1), 3-12 (2000)
MEDLINE	20584711
REFERENCE	3 (bases 1 to 990)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : seque@genoscope.cns.fr - Web : www.genoscope.cns.fr) This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 kb 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
FEATURES	Location/Qualifiers 1..990 /organism="Kluyveromyces thermotolerans" /strain="CBS 6340" /db_xref="taxon:4916" /collection="AY0AA007FE09" /clone.lib="AY0AA" /note="end : 17" <241..>471 /note="similar to P78771 [Schizosaccharomyces pombe, unknown protein]" /evidence=not_experimental
SOURCE	BASE COUNT 219 a 272 c 265 g 231 t 3 others
ORIGIN	
misc-feature	
Query Match	55.3%; Score 21; DB 13;
Best Local Similarity	73.0%; Pred. NO. 2.3e+02;
Matches 27; Conservative	0; Mismatches 10; Indels 0; Gaps 0.
Gy	1 tgcttagaggaatccgaagtgcagatacttgcg 37 Db 176 tcggccaggcgctctgaagtcaccgctttggcg 140
RESULT 10	BG619393 734 bp mRNA EST 18-APR-2001
LOCUS	BG619393/C 60261913871 NIH_MGC_79 Homo sapiens CDNA clone IMAGE:4732929 5',
DEFINITION	mRNA sequence.
ACCESSION	BG619393
VERSION	BG619393.1 GI:13670764
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 734)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgapubs@email.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.lnl.nih.gov>
Plate: LNCMI593 row: k column: 10
High quality sequence stop: 701.

FEATURES source location/Qualifiers
1..734
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image=4732929"
/clone_id="NIH_MGC_79"
/_lab_host="DH10B (TI phage-resistant)"
/note="Organ: placenta; Vector: pDNR-LTB (Clontech); Site-1: SfiI (ggccgctcgccc); Site-2: SfiI (ggcattatgcc) : 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCGTAGAGGCCGACGGCGCCGCACTG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.3 kb (range 0.5-4.0 kb), 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

BASE COUNT 223 a 139 c 127 g 245 t
ORIGIN

Query Match 54.7%; Score 20.8; DB 11; Length 734;
Best Local Similarity 78.1%; Pred. No. 2.6e+02;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 tgctctagagaagtccaagtacatacttt 32
||||| ||||| | ||||| | |||||
Db 477 TGCTTAATAAATTGGAAGTTAACCTPATTT 446

RESULT 11
AO696864/c DNA GSS 06-JUL-1999
LOCUS HS_5526_AZ_D03_SPG6 RPCI-11 Human Male BAC Library Homo sapiens
DEFINITION genomic clone Plate-1102 Col-6 Row-G, DNA sequence.
ACCESSION AO696864
VERSION AO696864.1 GI:5387112
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 430)
AUTHORS Mahairas,G.C., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Kelder,A., Shaker,R., Furlong,J.J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
CONTACT: Mahairas GC, Wallace JC, Hood L
HIGH THROUGHPUT SEQUENCING CENTER
UNIVERSITY OF WASHINGTON
401 Queen Anne Avenue North, Seattle, WA 98109, USA
TEL: (206) 616-3618
FAX: (206) 616-3887
EMAIL: jwallace@u.washington.edu
CLONES ARE DERIVED FROM THE HUMAN BAC LIBRARY RPCI-11. FOR BAC

KEYWORDS
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
TITLE Strausberg, R.
AUTHORS Direct Submission
JOURNAL Submitted (12-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-ref@mail.nih.gov
Tissue Procurement: ANCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland:
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nigr.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McQuinn, J., Pearson, R., Snyder, B., Stantipop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgou, C., Vogt, J.L., Walker, M.A., Zhang, L., H. and Green, E.D.

FEATURES
source
1. .906
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:412154"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_id="NHL_MGC_17"
/lab_host="DH10B-R"
/note="Vector: pORF7"

BASE COUNT
317 a 170 c 307 g 112 t

ORIGIN

Query Match 57.9%; Score 22; DB 12; Length 906;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 17 gaagtgaacatacttggcgc 38
|||||

Db 12 GAAGTGAACATATCTTGGCGC 33

RESULT 5
LOCUS A1728511 564 bp mRNA EST 11-JUN-1999
DEFINITION BNGH110926 Six-day cotton fiber Gossypium hirsutum cDNA 5' similar to FERROCHELATASE PRECURSOR (PROHEME FERRO-LYASE) (HME SYNTHETASE) g114749681gnl|PID|d1005644 (D26106) ferrochelatase [Gossypium sativum], mRNA sequence.

ACCESSION
VERSION A1728511
KEYWORDS EST.
SOURCE upland cotton.
ORGANISM Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE
AUTHORS Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
TITLE Biewitt, M., Matz, E.C., Davy, D.F. and Burr, B.
JOURNAL ESTs from developing cotton fiber
COMMENT Unpublished (1999)
Contact: Ben Burr
Brooklyn Department
Brooklyn National Laboratory
Upton, NY 11973, USA
Tel: 516-344-3396
Fax: 516-344-3407
Email: burr@nslu1.bnl.gov
Seq primer: T3 Primer.

FEATURES
source
1. .564
/organism="Gossypium hirsutum"
/cultivar="Acala Maxxa"
/db_xref="taxon:3635"
/clone_lib="Six-day Cotton fiber"
/tissue_type="Immature fiber"
/dev_stage="Six days post anthesis"
/lab_host="XLI-Blue"
/note="Vector: pBluescript II KS+"

BASE COUNT
152 a 107 c 144 g 160 t 1 others

ORIGIN

Query Match 56.3%; Score 21.4; DB 10; Length 564;
Best Local Similarity 80.6%; Pred. No. 1.4e+02;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 4 tctagaagatcgaatgacatacttg 34
|||||

Db 273 TCTAGAGATTCGAATGATCAACCTTG 303

RESULT 6
LOCUS A2768917/c 498 bp DNA GSS 16-FEB-2001
DEFINITION 1M0569E11F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0569E11 F, DNA sequence.

ACCESSION
VERSION A2768917
KEYWORDS A2768917.1 GI:12888508
GSS.

KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A. and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10Kb plasmid inserts
JOURNAL Unpublished (2000)
CONTACT: Robert B. Weiss
UNIVERSITY OF UTAH
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0569 row: E column: 11
Seq primer: CGTGTAAACGACGCGCAGT
CLASS: plasmid ends
High quality sequence stop: 498.

FEATURES
source
1. .498
/organism="Mus musculus"
/strain="C57BL/6J"

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:14:26 ; Search time 3881.49 Seconds
(without alignments)
105.202 Million cell updates/sec

Title: US-09-915-060-7
Perfect score: 38
Sequence: 1 tgcctcagaggaattcgaagtgcagctactttggcgc 38

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: em_estfun:**
- 2: em_esthum:**
- 3: em_estin:**
- 4: em_estom:**
- 5: em_estopl:**
- 6: em_estba:**
- 7: em_estro:**
- 8: em_estov:**
- 9: em_hlc:**
- 10: gb_est1:**
- 11: gb_est2:**
- 12: gb_hlc:**
- 13: gb_gss:**
- 14: em_gss_fun:**
- 15: em_gss_hum:**
- 16: em_gss_inv:**
- 17: em_gss_pln:**
- 18: em_gss_pro:**
- 19: em_gss_rnd:**
- 20: em_gss_vrt:**
- 21: em_gss_other:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22.2	58.4	411	AA159840	AA159840 z056g07.r
2	22.2	57.9	488	AA213378	AA213378 z056g07.s
3	22.2	57.9	640	11 BG827010	BG827010 602749042
4	22.2	57.9	906	12 BC009375	BC009375 Homo sapi
5	21.4	56.3	564	10 A1728511	A1728511 BMTG1109
6	21.4	55.3	498	13 A2768917	A2768917 IM0569E11
7	21.4	55.3	618	10 AU094789	AU094789 AU094789
8	21.4	55.3	763	10 A1574604	A1574604 A1574604
9	21.4	55.3	990	11 CNS06YA2	CNS06YA2
10	20.6	54.7	734	11 BG619393	BG619393 602619138
11	20.6	54.2	430	13 A0696864	A0696864 HS_5526_A
12	20.6	54.1	13	A0511435	A0511435 HS_5214_A

C 13	20.6	54.2	642	10	AT638929	AT638929 AEMTAL50
C 14	20.6	54.2	1117	13	CNS07C7R	AT438829 T3 end of
C 15	20.4	53.7	546	13	A0402410	HS_5060_A
C 16	20.4	53.2	300	11	C07641	C07641 C07641 Y0J1
C 17	20.2	53.2	352	11	R03855	R03855 PK12609.s1
C 18	20.2	53.2	660	13	A2471544	A2471544 1M0286011
C 19	20.2	53.2	666	13	A0545121	CITB1-E1
C 20	20.2	53.2	822	10	AM9833246	HVSMERG00
C 21	20.2	53.2	881	11	BF244937	601864340
C 22	20.2	52.6	306	10	A1524605	t043h04.x
C 23	20.2	52.6	327	10	AA663181	ab74c01.s
C 24	20.2	52.6	432	11	BF566267	UI-R-BT1.s
C 25	20.2	52.6	491	10	BE511204	BE511204 946058F05
C 26	20.2	52.6	510	11	BG411126	EM1_26_A0
C 27	20.2	52.6	516	13	A0721424	HS_5559_A
C 28	20.2	52.6	613	10	AM455707	707090809
C 29	20.2	52.6	684	13	CNS01S1S	AL165673 Tetraodon
C 30	20.2	52.6	798	11	BG321018	zm04_02h0
C 31	20.2	52.6	892	11	BG321059	zm04_02h0
C 32	20.2	52.6	1403	11	BG114146	BG114146 602285564
C 33	19.8	52.1	140	10	AA840946	MB3D6A1F
C 34	19.8	52.1	414	13	A0145635	HS_2218_B
C 35	19.8	52.1	525	13	TA293601P	T. brucei
C 36	19.8	52.1	527	11	BG447599	NF117B10S
C 37	19.8	52.1	570	10	A1438619	486012F06
C 38	19.8	52.1	616	10	AM687972	NF001F06S
C 39	19.8	52.1	636	10	AM689592	NF022B03S
C 40	19.8	52.1	640	10	A1438623	486012S01
C 41	19.8	52.1	654	10	AM695839	NF099B08S
C 42	19.8	52.1	664	10	BE204210	EST396886
C 43	19.8	52.1	719	11	BF641446	NF056D11I
C 44	19.8	52.1	719	11	BG645842	EST507461
C 45	19.6	51.6	300	11	C59696	C59696 C59696 Y0J1

ALIGNMENTS

RESULT 1
AA159840/c
LOCUS
DEFINITION
IMAGE:590940 5' similar to SW:DBR1_YEAST P24309 LARIAT DEBRANCHING
ENZYME: mRNA sequence.

ACCESSION
AA159840
VERSION
AA159840.1 GI:1734985
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiappelli,B., Chiusso,S., Dietrich,N., Dubuque,T., Favello,A., Gish,M., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Madis,E., Moore,B., Morris,M., Parsons,D., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,D., Trevisan,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Mair,M.

Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 330.

FEATURES

source
1. 411

;; FILING DATE: 7-JUNE-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Barry J. Swanson
;; REGISTRATION NUMBER: 33,215
;; REFERENCE/DOCKET NUMBER: NEX66
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (303) 268-0066
;; TELEFAX: (303) 268-0065
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 86 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA
;; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-08-991-743C-14

Query Match 47.9%; Score 18.2; DB 4; Length 86;
Best Local Similarity 74.2%; Pred. No. 26;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tgctctagaggaattcgaaatgacgatactt 31
Db 70 TGCTCAAGTGGGACTGAGAGTGACGAGGCAT 40

RESULT 15
PCT-US96-08014-103/c
; Sequence 103, Application PC/TUS9608014
; GENERAL INFORMATION:
; APPLICANT: LARRY GOLD; NEBOUSA JANITC; STEVEN RINGQUIST; NIKOS
; TITLE OF INVENTION: HIGH AFFINITY OLIGONUCLEOTIDE
; TITLE OF INVENTION: LIGANDS TO TRANSFORMING GROWTH
; TITLE OF INVENTION: FACTOR TO (TGF), PLATELET-DERIVED
; TITLE OF INVENTION: GROWTH FACTOR (PDGF) AND HUMAN
; NUMBER OF INVENTION: KEROTINOCYTE GROWTH FACTOR (hKGF)
; NUMBER OF SEQUENCES: 304
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; City: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
; COMPUTER: IBM compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/08014
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/458,423
; FILING DATE: 02-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/458,424
; FILING DATE: 02-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/465,594
; FILING DATE: 05-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/465,591
; FILING DATE: 05-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,725
; FILING DATE: 07-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/479,783

;; FILING DATE: 07-JUNE-1995
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/618,693
;; FILING DATE: 20-MARCH-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Barry J. Swanson
;; REGISTRATION NUMBER: 33,215
;; REFERENCE/DOCKET NUMBER:
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (303) 793-3333
;; TELEFAX: (303) 793-3333
;; INFORMATION FOR SEQ ID NO: 103:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 86 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA
PCT-US96-08014-103

Query Match 47.9%; Score 18.2; DB 5; Length 86;
Best Local Similarity 74.2%; Pred. No. 26;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 tgctctagaggaattcgaaatgacgatactt 31
Db 70 TGCTCAAGTGGGACTGAGAGTGACGAGGCAT 40

Search completed: December 4, 2001, 18:58:38
Job time: 9264 sec


```

: TELEPHONE: (703)836-9300
: TELEFAX: (703)663-4109
:
: TELE: 899149
:
: INFORMATION FOR SEQ ID NO: 31:
:
: SEQUENCE CHARACTERISTICS:
:
: LENGTH: 11298 base pairs
:
: TYPE: NUCLEIC ACID
:
: STRANDEDNESS: single
:
: TOPOLOGY: linear
:
: ORIGINAL SOURCE:
:
: ORGANISM: homo sapien
:
: STRAIN: FCRI beta
:
: OS-07-869-933-31

```

Query Match	48.9%	Score 18.6	DB 1	Length 11298
Best Local Similarity	72.7%	Pred. No. 49		
Matches 24	Conservative 0	Mismatches 9	Indels 0	Gaps 0

```
OY      1  tgcctagagaattcgaagtgaacatacttt 33
        |||||
Db    5630 TGTTCATAGGAATTGTAAGTATGATGCCTT 5662
```

RESULT 5
US-08-201-879A-2

1 GENERAL INFORMATION:
2 APPLICANT: KINET, Jean-Pierre
3 APPLICANT: JOUVIN, Marie-Helene
4 TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
5 TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
6 TITLE OF INVENTION: IMMUNOGLOBULIN E
7 NUMBER OF SEQUENCES: 5
8 CORRESPONDENCE ADDRESS:
9 ADDRESSEE: Foley & Lardner
10 STREET: 3000 K Street, N.W., Suite 500
11 CITY: Washington
12 STATE: D.C.

```

; STRAIN: FCRI beta
;
; FEATURE:
;
; NAME/KEY: CDS
; LOCATION: join(456..511, 1381..1510, 2026..2160, 4475..4531)
; LOCATION: 5079..5237, 5640..5738, 7224..7319)
;
US-08-201-879A-2

```

Query Match	48.9%	Score 18.6	DB 1	length 11298
Best Local Similarity	72.7%	Pred No. 49		
Matches 24	Conservative 0	Mismatches 9	Indels 0	Gaps 0

```
QY      1  tgcctagaggaaatccgaagtgcacatacltt 33
          || || | || || | || || || || ||
Db    5630  TGTTCAATAGGAATTGTAGTGATGATGCCTT 5662
```

RESULT 6
US-09-103-663-31

```

: Sequence 31, Application US/09103663D
: Patent No. 6171803
:
: GENERAL INFORMATION:
:
: APPLICANT: Kinet et al.
: TITLE OF INVENTION: Isolation, characterization, and use of the human beta
: TITLE OF INVENTION: subunit of the high affinity receptor for
: TITLE OF INVENTION: immunoglobulin E.
: FILE REFERENCE: 50490
: CURRENT APPLICATION NUMBER: US/09/103,663D
: CURRENT FILING DATE: 1998-06-23
: EARLIER APPLICATION NUMBER: 07/869,933
: EARLIER FILING DATE: 1992-04-16
: NUMBER OF SEQ ID NOS: 35
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 31
:
: LENGTH: 11298
:
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-103-663-31

```

Query Match	48.9%	Score 18.6	DB 4	Length 11298
Best Local Similarity	72.7%	Pred. No. 49		
Matches 24	Conservative	0	Mismatches 9	Indels 0
				Gaps 0

QY 1 tgcctagagaattcgaagttagcgatactttt 33
|| || | ||||| | | ||||| || | ||
Db 5630 tggtcaataggaatatgtagtgatgatgctgtt 5662

RESULT 7
US-08-249-013-1
; Sequence 1, Application US/08249013

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Db 304 GCTCTGCAGGATTTTCATGTCACCATCCTTTGCC 270

RESULT 2
 US-08-458-356-163/c
 Sequence 163, Application US/08458356
 Patent No. 5942235
 GENERAL INFORMATION:
 APPLICANT: Paoletti, Enzo
 APPLICANT: Tantiaglia, James
 APPLICANT: Cox, William I.
 TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
 NUMBER OF SEQUENCES: 217
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Curtis, Morris & Safford
 STREET: 530 Fifth Avenue
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/458,356
 FILING DATE: 02-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/164,009
 FILING DATE: 19-JAN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Frommer, William S.
 REGISTRATION NUMBER: 25,506
 REFERENCE/DOCKET NUMBER: 454310-2530
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 840-3333
 TELEFAX: (212) 840-0712
 TELEX: 425066CURTWS
 INFORMATION FOR SEQ ID NO: 163:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 598 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 IS-08-458-356-163

Query Match	50.0%	Score 19;	DB 2;	Length 598;
Best Local Similarity	71.4%	Pred. No. 17;		
Matches	25; Conservative	0; Mismatches	10; Indels	0; Gaps
OY	2 gctctagagaattcgaagtacgataactttggc	36		
Ib				
	304 GCCTGCACAGATTTCATGTCACCATCCTTTGGCC	270		

RESULT 3
 US-08-460-736-163/c
 : Sequence 163 Application US/08460736
 : Patent No. 6265189
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Paoletti, Enzo
 : APPLICANT: Tartaglia, James
 : APPLICANT: Cox, William I.
 : TITLE OF INVENTION: RECOMBINANT VIRUS IMMUNOTHERAPY
 : NUMBER OF SEQUENCES: 217
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Curtiss, Morris & Safford
 : STREET: 530 Fifth Avenue
 : CITY: New York

```

1 STATE: NY
2 COUNTRY: USA
3 ZIP: 10036
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: Floppy disk
7 COMPUTER: IBM PC compatible
8 OPERATING SYSTEM: PC-DOS/MS-DOS
9 SOFTWARE: PatentIn Release #1.0, Version #1.25
10
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/08/460,736
13 FILING DATE: 02-JUN-1995
14
15 CLASSIFICATION: 514
16
17 PRIOR APPLICATION DATA:
18 APPLICATION NUMBER: US 08/184,009
19 FILING DATE: 19-JAN-1994
20 ATTORNEY/AGENT INFORMATION:
21 NAME: Frommer, William S.
22 REGISTRATION NUMBER: 25,506
23 REFERENCE/DOCKET NUMBER: 454310-2530
24
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: (212) 840-3333
27 TELEFAX: (212) 840-0712
28
29 TELEX: 425066CGRFMS
30
31 INFORMATION FOR SEQ ID NO: 163:
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 598 base pairs
34 TYPE: nucleic acid
35 STRANDEDNESS: single
36 TOPOLOGY: linear
37
38 MOLECULE TYPE: CDNA
39
40 US-08-460-736-163

```

Query Match	50.0%;	Score 19;	DB 4;	Length 598;
Best Local Similarity	71.4%;	Pred. No. 17;		
Matches 25;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;

RESULT 4
US-07-869-933-31
; Sequence 31, Application US/07869933
; Patent No. 5770396
; GENERAL INFORMATION:
; APPLICANT: KINER, Jean-Pierre
; TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
; TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
; TITLE OF INVENTION: IMMUNOGLOBULIN
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/869,933
; FILING DATE: 19920416
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/154 NIHD
; TELECOMMUNICATION INFORMATION:

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:58:37 ; Search time 152.61 Seconds (without comments)

Title: US-09-915-060-7

Sequence: 1 **tgctctagaggaattcgaagtcacgatactt**tgcgcg 38

Scoring table: IDENTITY_NUC

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

```
Database : Issued_Patents_NA: *
1. /usr36/databases/0/ins/5A COMP sec.*
```

```

1: /cgn2_6/ptodata_2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata_2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata_2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata_2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata_2/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata_2/ina/backfiles1.seq:

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	19	50.0	598	2	US-08-184-009-163	Sequence 163, App1
C 2	19	50.0	598	2	US-08-458-556-163	Sequence 163, App1
C 3	19	50.0	598	4	US-08-460-736-163	Sequence 163, App1
C 4	18.6	48.9	11298	1	US-07-869-933-31	Sequence 31, App1
5	18.6	48.9	11298	1	US-08-201-879A-2	Sequence 2, App1
6	18.6	48.9	11298	1	US-09-103-663-31	Sequence 31, App1
7	18.4	48.4	1991	1	US-08-249-013-1	Sequence 1, App1
8	18.4	48.4	1991	1	US-08-886-863-1	Sequence 1, App1
9	18.4	48.4	1991	5	PCT-US95-06764-1	Sequence 1, App1
C 10	18.2	47.9	86	1	US-08-479-783A-14	Sequence 14, App1
C 11	18.2	47.9	86	1	US-08-479-725-14	Sequence 14, App1
C 12	18.2	47.9	86	1	US-08-618-693-14	Sequence 14, App1
C 13	18.2	47.9	86	4	US-08-973-124-103	Sequence 103, App1
C 14	18.2	47.9	86	4	US-08-991-743C-14	Sequence 14, App1
C 15	18.2	47.9	86	5	PCT-US96-08014-103	Sequence 103, App1
16	18	47.4	2236	2	US-08-484-993B-13	Sequence 13, App1
17	18	47.4	2236	2	US-08-484-158B-13	Sequence 13, App1
18	18	47.4	2236	2	US-08-484-596A-13	Sequence 13, App1
19	18	47.4	2236	2	US-08-480-150A-13	Sequence 13, App1
20	18	47.4	2236	3	US-08-148-731-13	Sequence 13, App1
21	18	47.4	2236	3	US-08-159-223A-13	Sequence 13, App1
22	17.8	46.8	3205	4	US-09-061-768A-3	Sequence 3, App1
23	17.8	46.8	4060	1	US-08-164-292B-1	Sequence 1, App1
24	17.8	46.8	4060	1	US-08-164-292B-3	Sequence 3, App1
25	17.8	46.8	4060	1	US-08-164-292B-5	Sequence 5, App1
26	17.8	46.8	4060	1	US-08-164-292B-7	Sequence 7, App1
27	17.8	46.8	4060	3	US-08-845-623-31	Sequence 1, App1

28	17.8	46.8	4060	3	US-08-845-622-3	Sequence 3, April
29	17.8	46.8	4060	3	US-08-845-623-5	Sequence 5, April
30	17.8	46.8	4060	3	US-08-845-623-7	Sequence 7, April
31	17.8	46.8	4060	3	US-08-815-927-1	Sequence 1, April
32	17.8	46.8	4060	3	US-08-815-927-3	Sequence 3, April
33	17.8	46.8	4060	3	US-08-815-927-5	Sequence 5, April
34	17.8	46.8	4060	3	US-08-815-927-7	Sequence 7, April
35	17.6	46.3	305	1	US-08-170-294-5	Sequence 5, April
36	17.6	46.3	305	1	US-08-664-855-5	Sequence 5, April
37	17.6	46.3	305	1	US-09-049-288-5	Sequence 5, April
38	17.6	46.3	2042	2	US-08-933-821-16	Sequence 16, April
39	17.6	46.3	2042	3	US-08-933-821-16	Sequence 16, April
40	17.6	46.3	2042	3	US-08-960-507-16	Sequence 16, April
41	17.6	46.3	2042	3	US-09-143-086-5	Sequence 5, April
42	17.6	46.3	2493	3	US-08-845-055-3	Sequence 3, April
43	17.6	46.3	3827	1	US-08-170-294-6	Sequence 6, April
44	17.6	46.3	3827	2	US-08-664-855-6	Sequence 6, April
45	17.6	46.3	3827	2	US-08-718-751-1	Sequence 1, April

ALIGNMENTS

RESULT 1
HS-08-184-000-163/0

; Sequence 163, Application US/08184005
Patent No. 5933075

; GENERAL INFORMATION:

APPLICANT: Tartaglia, James
ADDRESSEE: Coy William T

TITLE OF INVENTION:	RECOMBINANT VIRUS IMMUNOTHERAPY
NUMBER OF SENTENCES:	217

CORRESPONDENCE ADDRESS:

STREET: 530 Fifth Avenue

STATE: NY

```

; ZIP: 10036

```

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 100 400 000

```

FILING DATE: 19-JAN-1994

ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 25,506

TELECOMMUNICATION INFORMATION
SERIES 1-800-368-5868

TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 163
SEQUENCE CHARACTERISTICS:

LENGTH: 598 base pairs

```
STRANDEDNESS: single
```

MOLECULE TYPE: CDNA
;
HE-09-184-009-163

Query Match	50.0%;	Score 19;	DB 2;	Length 598
-------------	--------	-----------	-------	------------

Matches	25;	Conservative	0;	Mismatches	10;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	-----	--------	----	------	----

2 gctctagaggaattcgaagtgcacatacttctgc 36


```

PI Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
PI White OR;
XX
XX WPI: 1999-081217/07.
XX
XX New isolated Borrelia burgdorferi nucleic acids - used to develop
XX products for the detection, diagnosis, characterisation, prevention
XX and therapy of infections, particularly Lyme disease
XX
XX Claim 1: Page 831-851; 1128pp; English.
XX
XX AA020248 to AA020402 represent polynucleotide sequences isolated from
XX Borrelia burgdorferi (Bb). Products derived from Bb can be used for
XX the detection, diagnosis, characterisation, prevention and therapy of
XX Bb infections, e.g. Lyme disease. They can also be used for the
XX production of biosynthetic products, e.g. enzymes. Borrelia belongs
XX to a family of motile, spiral-shaped bacteria called Spirochetes.
XX Spirochetes are pathogenic in humans and Borrelia causes epidemic and
XX endemic relapsing fever, and Lyme borreliosis, more commonly known as
XX Lyme disease.
XX
XX Sequence 35515 BP; 10795 A; 5947 C; 4439 G; 13907 T; 427 other;
XX
XX
XX Query Match
XX Best Local Similarity 50.5%; Score 19.2; DB 20; Length 35515;
XX Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
XX
XX
XX 3 cctagaggaattcgaagtcagatcttgg 34
XX ||||| ||| ||||| ||||| ||
XX 7225 CTCCTAGAGTAACATTAGTGAAGATCTTTAG 7194
XX
XX
XX RESULT 13
XX AA022774/C
XX ID AA022774 standard; cDNA to mRNA; 151 BP.
XX
XX AC AA022774;
XX
XX DT 15-AUG-1996 (first entry)
XX
XX DE Human gene signature HUMGS04434.
XX
XX KW Gene signature: messenger RNA; mRNA; relative abundance; frequency;
XX human; cloning; mapping; non-biased library; diagnosis; detection;
XX cell typing; abnormal cell function; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO9514772-A1.
XX
XX PD 01-JUN-1995.
XX
XX PF 11-NOV-1994; 94WO-JP01916.
XX
XX PR 12-NOV-1993; 93JP-0355504.
XX
XX PA (MATSU) MATSUBARA K.
XX PA (OKUBO) OKUBO K.
XX
XX PI Matsubara K, Okubo K;
XX
XX DR WPI: 1995-206931/27.
XX
XX PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
XX for diagnosis of abnormal cell function, by preparing cDNA that
XX reflects relative abundance of corresp. mRNA in specific human
XX tissues
XX
XX PS Claim 1: Page 1211; 2245pp; Japanese.
XX
XX CC A single-stranded DNA (or its complementary strand or the corresp.
XX double-stranded DNA) which comprises one of the 7837 "GS" sequences
XX
XX LC

```

```

CC given in AAT19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
XX
XX Sequence 151 BP; 58 A; 19 C; 24 G; 50 T; 0 other;
XX
XX
XX Query Match
XX Best Local Similarity 50.0%; Score 19; DB 16; Length 151;
XX Matches 25; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
XX
XX
XX 2 gctctagaggaattcgaagtcagatcttggc 36
XX ||||| ||||| ||| ||| ||| |||
XX 41 GCTCAAGAGGAATTCAGATTACTGTTTATGTC 7
XX
XX
XX RESULT 14
XX AA067872/C
XX ID AA067872 standard; DNA; 598 BP.
XX
XX AC AA067872;
XX
XX DT 23-MAR-1995 (first entry)
XX
XX DE IL3/Murine IFN-gamma expression cassette.
XX
XX KW Polymerase chain reaction; primer; amplify; NVVAC; ALVAC; recombinant;
XX murine; interleukin-2; IL-2; PRM825; pmu-1; PBS-SK; pKM151; TK vector;
XX plasmid; vaccinia; H6 promoter; amplify; primer; antigenic response;
XX polymerase chain reaction; poxvirus; PSD542; immunological response;
XX pathogen; ss.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT CDS 101..598
XX FT
XX FT promoter
XX FT
XX FT
XX FT
XX PN WO9416716-A.
XX
XX PD 04-AUG-1994.
XX
XX PF 21-JAN-1994; 94WO-US00888.
XX
XX PR 21-JAN-1993; 93US-0007115.
XX PR 19-JAN-1994; 94US-0184009.
XX
XX PA (VIRO-) VIROGENETICS CORP.
XX
XX PI Cox WL, Paoletti E, Tartaglia J;
XX
XX DR WPI: 1994-263767/32.
XX
XX PT Attenuated recombinant virus used for cancer therapy - comprises
XX DNA encoding cytokine and/or tumour associated antigen
XX
XX PS Example 20: Fig 26; 232pp; English.
XX
XX CC This sequence represents an expression cassette containing the vaccinia

```

```
KW Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
KW infection; diagnosis; characterisation; detection; ds.
OS Borrelia burgdorferi.
XX
XX WO9858943-A1.
XX
XX 30-DEC-1998.
XX
XX 18-JUN-1998; 98WO-US12764.
XX
XX 03-SEP-1997; 97US-0057483.
XX 20-JUN-1997; 97US-0050359.
XX 22-JUL-1997; 97US-0053344.
XX 22-JUL-1997; 97US-0053377.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (MEDI-) MEDIMUNE INC.
XX
XX Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
XX White OR;
XX
XX WPI: 1999-081217/07.
XX
XX New isolated Borrelia burgdorferi nucleic acids - used to develop
XX products for the detection, diagnosis, characterisation, prevention
XX and therapy of infections, particularly Lyme disease
XX
XX Claim 1: Page 1023-1024; 1128pp; English.
XX
XX AAX20248 to AAX20402 represent polynucleotide sequences isolated from
XX Borrelia burgdorferi (Bb). Products derived from Bb can be used for
XX the detection, diagnosis, characterisation, prevention and therapy of
XX Bb infections, e.g. Lyme disease. They can also be used for the
XX production of biosynthetic products, e.g. enzymes. Borrelia belongs
XX to a family of motile, spiral-shaped bacteria called Spirochetes.
XX Spirochetes are pathogenic in humans and Borrelia causes epidemic and
XX endemic relapsing fever, and Lyme borreliosis, more commonly known as
XX Lyme disease.
XX
XX Sequence 2532 BP; 874 A; 451 C; 250 G; 955 T; 2 other;
XX
SQ
Query Match 50.5%; Score 19.2; DB 20; Length 2532;
Best Local Similarity 75.0%; Pred. No. 70;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 3 ctcagaggaattcgaagtgcatactttg 34
| | | | | | | | | | | | | | | | | |
DB 178 CTCTAGAGCTAACATTTAGTGAAGATTCTTTAG 147

RESULT 11
AAX20268/c
ID AAX20268 standard; DNA: 5805 BP.
XX
XX AAX20268;
XX
XX 04-MAY-1999 (first entry)
XX
XX Borrelia burgdorferi polynucleotide sequence #21.
XX
XX Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
XX epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
XX infection; diagnosis; characterisation; detection; ds.
XX
XX Borrelia burgdorferi.
XX
XX WO9858943-A1.
XX
XX 30-DEC-1998.
XX
```

```
PF 18-JUN-1998; 98WO-US12764.
XX
XX 03-SEP-1997; 97US-0057483.
XX 20-JUN-1997; 97US-0050359.
XX 22-JUL-1997; 97US-0053344.
XX 22-JUL-1997; 97US-0053377.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (MEDI-) MEDIMUNE INC.
XX
XX Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
XX White OR;
XX
XX WPI: 1999-081217/07.
XX
XX New isolated Borrelia burgdorferi nucleic acids - used to develop
XX products for the detection, diagnosis, characterisation, prevention
XX and therapy of infections, particularly Lyme disease
XX
XX Claim 1: Page 959-962; 1128pp; English.
XX
XX AAX20248 to AAX20402 represent polynucleotide sequences isolated from
XX Borrelia burgdorferi (Bb). Products derived from Bb can be used for
XX the detection, diagnosis, characterisation, prevention and therapy of
XX Bb infections, e.g. Lyme disease. They can also be used for the
XX production of biosynthetic products, e.g. enzymes. Borrelia belongs
XX to a family of motile, spiral-shaped bacteria called Spirochetes.
XX Spirochetes are pathogenic in humans and Borrelia causes epidemic and
XX endemic relapsing fever, and Lyme borreliosis, more commonly known as
XX Lyme disease.
XX
XX Sequence 5805 BP; 1878 A; 1000 C; 614 G; 2290 T; 23 other;
XX
SQ
Query Match 50.5%; Score 19.2; DB 20; Length 5805;
Best Local Similarity 75.0%; Pred. No. 81;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 3 ctcagaggaattcgaagtgcatactttg 34
| | | | | | | | | | | | | | | | | |
DB 1242 CTCTAGAGCTAACATTTAGTGAAGATTCTTTTG 1211

RESULT 12
AAX20252/c
ID AAX20252 standard; DNA: 35515 BP.
XX
XX AAX20252;
XX
XX 04-MAY-1999 (first entry)
XX
XX Borrelia burgdorferi polynucleotide sequence #5.
XX
XX Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
XX epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
XX infection; diagnosis; characterisation; detection; ds.
XX
XX Borrelia burgdorferi.
XX
XX WO9858943-A1.
XX
XX 30-DEC-1998.
XX
XX 18-JUN-1998; 98WO-US12764.
XX
XX 03-SEP-1997; 97US-0057483.
XX 20-JUN-1997; 97US-0050359.
XX 22-JUL-1997; 97US-0053344.
XX 22-JUL-1997; 97US-0053377.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (MEDI-) MEDIMUNE INC.
XX
```


AC AAC93671;
 XX
 DT 19-FEB-2001 (first entry)
 XX
 DE Cat flea HNC CLB2A cDNA complement, SEQ ID NO:164.
 XX
 KM Cat flea: head and nerve cord nucleic acid; HNC: flea infestation;
 XX vaccine: antiparasitic; therapeutic target; diagnosis; detection; ss.
 XX
 OS Ctenocephalides felis.
 XX
 PN W0200061621-A2.
 XX
 PD 19-OCT-2000.
 XX
 PF 07-APR-2000; 2000MO-US09437.
 XX
 PR 09-APR-1999; 99US-0128704.
 XX
 PA (HESK-) HESKA CORP.
 XX
 PI Brandt KS, Gaines PJ, Stinchcomb DR, Wisniewski N;
 XX P-PSDB; AAB29618.
 DR WPI: 2000-656323/63.
 XX
 PT Flea Malpighian tubule and head and nerve cord tissue derived nucleic
 PT acids useful for the prevention, diagnosis and treatment of flea
 PT infestations -
 XX
 PS Claim 1: Page 298; 964pp: English.
 XX
 CC The invention relates to novel cat flea (Ctenocephalides felis) nucleic
 CC acids which are expressed in hindgut and Malpighian tubule (HMT) tissue
 CC or head and nerve cord (HNC) tissue. The invention also relates to the
 CC encoded proteins. The invention additionally encompasses expression
 CC constructs, recombinant viruses and recombinant cells comprising the
 CC nucleic acids of the invention, recombinant production of the proteins,
 CC antibodies against the proteins, a method of identifying inhibitors of
 CC the proteins, and compositions comprising the inhibitors for
 CC administration to an animal. The nucleic acids, and the proteins they
 CC encode may be used in the prevention, treatment and diagnosis of diseases
 CC associated with flea infestations. For example, the nucleic acids may be
 CC used to produce an HMT or HNC protein according to standard recombinant
 CC DNA methodology by inserting the nucleic acids into a host cell and
 CC culturing the cell to express the protein. The HMT and HNC nucleic acids
 CC may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect
 CC and quantitate the presence of cat flea or other homologous nucleic acid
 CC sequences in samples. They may also be used to study the expression and
 CC function of the proteins and their role in metabolism. The HMT and HNC
 CC proteins may be used as antigens in the production of specific
 CC antibodies, and in assays to identify modulators (agonists and
 CC antagonists) of HMT and/or HNC protein expression and activity. The
 CC anti-HMT/HNC protein antibodies and antagonists may also be used to
 CC downregulate protein expression and activity. The antibodies may also be
 CC used as diagnostic agents for detecting the presence of flea polypeptides
 CC in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
 CC present sequence represents a cat flea HNC cDNA of the invention.
 XX
 SO Sequence 631 BP; 150 A; 144 C; 82 G; 255 T; 0 other;

Query Match 50.5%; Score 19.2; DB 21; Length 631;
 Best Local Similarity 75.0%; Pred. No. 55;
 Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 6 tagagaattcgaagtgaactactttggc 37
 ||||| ||| ||| ||| ||| ||| ||| |||
 DB 581 TACAGAGAAATCAAAATGATGATATTCTTGTCG 550

RESULT 9
 AAX13016/c

ID AAX13016 standard; DNA: 2213 BP.
 XX
 AC AAX13016;
 XX
 DT 19-MAR-1999 (first entry)
 XX
 DE Enterococcus faecalis genome contig SEQ ID NO:79.
 XX
 KM Enterococcus faecalis; contig; detection; Enterococcal infection;
 XX vaccine; attenuation; computer readable medium; ds.
 XX
 OS Enterococcus faecalis.
 XX
 PN W09850555-A2.
 XX
 PD 12-NOV-1998.
 XX
 PF 04-MAY-1998; 98MO-US08985.
 XX
 PR 14-NOV-1997; 97US-0066009.
 PR 06-MAY-1997; 97US-0044031.
 PR 16-MAY-1997; 97US-0046655.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Barash SC, Dillon PJ, Kunsch CA;
 XX P-PSDB; AAB29618.
 DR WPI: 1999-045171/04.
 XX
 PT New Isolated Enterococcus faecalis polynucleotides and polypeptides
 PT - used to develop products for the detection of Enterococcus and for
 PT use in vaccines for prevention or attenuation of Enterococcus
 PT infection.
 XX
 PS Claim 1: Page 558-559; 2084pp: English.
 XX
 CC A computer readable medium has been developed which has recorded on it
 CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
 CC AAX12938 to AAX13919 represent these nucleotide sequences which are
 CC primary nucleotide sequences, also known as contigs. The computer-based
 CC system can identify fragments of the Enterococcus faecalis genome with
 CC commercial importance. The products can be used to detect the presence
 CC of Enterococcus faecalis in samples. They can also be used for
 CC diagnosing Enterococcal infection in an animal and monitoring
 CC progression of disease, and for identifying agents which can be used to
 CC modulate the growth or pathogenicity of Enterococcus faecalis, or
 CC another related organism, in vivo or in vitro. In particular the
 CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
 CC can be used in vaccines to prevent or attenuate an Enterococcal
 CC infection.
 XX
 SO Sequence 2213 BP; 721 A; 342 C; 427 G; 720 T; 3 other;

Query Match 50.5%; Score 19.2; DB 20; Length 2213;
 Best Local Similarity 75.0%; Pred. No. 68;
 Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

OY 4 tctagaattcgaagtgaactactttgg 35
 ||||| ||| ||| ||| ||| ||| ||| |||
 DB 1008 TCTGAGCAATTGGAAATGGCGATTCTTGTTG 977

RESULT 10
 AAX20295/c
 ID AAX20295 standard; DNA: 2532 BP.
 XX
 AC AAX20295;
 XX
 DT 04-MAY-1999 (first entry)
 XX
 DE Borrelia burgdorferi polynucleotide sequence #48.
 XX

OY 4 tctagaagaattcgaatgcacatactttg 34
||||| ||
DB 308 TCTAGAGTAACAAATTACTGAAGACTTTTG 278

RESULT 6
AAA60982
ID AAA60982 standard; DNA: 1094 BP.

AC AAA60982;

DT 08-NOV-2000 (first entry)

DE Antibody nucleotide sequence.

KW Lida-protein; LDP; Lida-chromophore; LDC; antibody; lidamycin;
KM antibiotic; anticancer; tumour; cancer; ds.

OS Synthetic.

PN CNI251840-A.

PD 03-MAY-2000.

PF 13-OCT-1999; 99CN-0121668.

PR 13-OCT-1999; 99CN-0121668.

PA (MED1-) INST MEDICINAL BIOTECHNOLOGY CHINESE ACA.

PI Zhen Y, Li S, Jiang M;

DR WPI; 2000-432052/38.

PT Constitutive fusion protein of ridamycin as anticancer antibiotic and
single-chain antibody -

PS Disclosure; Page 2-3; 18pp; Chinese.

CC The present invention describes a fusion protein, designated lidamycin,
composed of lida-protein (LDP) and lida-chromophore (LDC), which acts as
an anticancer antibiotic and single-chain antibody. LDP and LDC are
joined via non-covalent bonds and can be splitted and recombined. A
single-chain antibody scfv can specifically join with IV-type collagenase
and can suppress its activity and in order to increase the permeability
of medicine to capillary tubes and the penetrability to real tumour, DNA
recombination and molecular recombination are used to prepare new-type
constitutive fusion protein LDM-FV of anticancer targeting medicine
lidamycin and single-chain antibody. Its molecular weight is about 37KDa.
CC It has the activity to inhibit IV-type collagenase and strong intrusion-
resistance kill action to cancer cells. It may be an ideal clinic
medicine. The present sequence represents an antibody nucleotide sequence
given in the exemplification of the present invention.

CC given in the exemplification of the present invention.

XX Sequence 1094 BP; 230 A; 312 C; 325 G; 227 T; 0 other;

SO Query Match 51.1%; Score 19.4; DB 21; Length 1094;

Best Local Similarity 70.3%; Pred. No. 50; Mismatches 11; Indels 0; Gaps 0;

OY 1 tgcctagaagaattcgaatgcacatactttg 37
||| ||| | ||||| ||||| |||||

DB 460 tctctgcagagccagtgaaagtgttcgatactatg 496

RESULT 7
AAC93670
ID AAC93670 standard; cDNA: 631 BP.
XX
AC AAC93670;
XX

DT 19-FEB-2001 (first entry)
XX
DE Cat flea HNC CLBP2A cDNA, SEQ ID NO:162.

KW Cat flea; head and nerve cord nucleic acid; HNC; flea infestation;
KM vaccine; antiparasitic; therapeutic target; diagnosis; detection; ss.

XX Ctenocephalides felis.

OS WO200061621-A2.

PN 19-OCT-2000.

PD 07-APR-2000; 2000WO-US09437.

PF 09-APR-1999; 99US-0128704.

PR (HESK-) HESKA CORP.

PA Brandt KS, Gaines PJ, Slinchcomb DT, Wisniewski N;

PI WPI; 2000-656323/63.

DR P-PsDB; AAB29618.

PT Flea Malpighian tubule and head and nerve cord tissue derived nucleic
acids useful for the prevention, diagnosis and treatment of flea
infestations -

PS Claim 1; Page 296-297; 964pp; English.

CC The invention relates to novel cat flea (Ctenocephalides felis) nucleic
acids which are expressed in hindgut and Malpighian tubule (HMT) tissue
or head and nerve cord (HNC) tissue. The invention also relates to the
CC encoded proteins. The invention additionally encompasses expression
CC constructs, recombinant viruses and recombinant cells comprising the
CC nucleic acids of the invention, recombinant production of the proteins,
CC antibodies against the proteins, a method of identifying inhibitors of
CC administration to an animal. The nucleic acids, and the proteins they
CC encode may be used in the prevention, treatment and diagnosis of diseases
CC associated with flea infestations. For example, the nucleic acids may be
CC used to produce an HMT or HNC protein according to standard recombinant
CC DNA methodology by inserting the nucleic acids into a host cell and
CC culturing the cell to express the protein. The HMT and HNC nucleic acids
CC may also be used as DNA probes in diagnostic assays (e.g., PCR) to detect
CC and quantitate the presence of cat flea or other homologous nucleic acid
CC sequences in samples. They may also be used to study the expression and
CC function of the proteins and their role in metabolism. The HMT and HNC
CC proteins may be used as antigens in the production of specific
CC antibodies, and in assays to identify modulators (agonists and
CC antagonists) of HMT and/or HNC protein expression and activity. The
CC anti-HMT/HNC protein antibodies and antagonists may also be used to
CC downregulate protein expression and activity. The antibodies may also be
CC used as diagnostic agents for detecting the presence of flea polypeptides
CC in samples (e.g., by enzyme linked immunosorbent assay (ELISA)). The
CC present sequence represents a cat flea HNC cDNA of the invention.

XX Sequence 631 BP; 255 A; 82 C; 144 G; 150 T; 0 other;

SO Query Match 50.5%; Score 19.2; DB 21; Length 631;

Best Local Similarity 75.0%; Pred. No. 55; Mismatches 8; Indels 0; Gaps 0;

OY 6 tagagaattcgaatgcacatactttg 37
||||| ||| ||| ||| ||| ||| |||

DB 51 tagaagaatacaaatgaagatacttctg 82

RESULT 8
AAC93671/c
ID AAC93671 standard; cDNA: 631 BP.
XX

```
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
```

```
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 52.6%; Score 20; DB 21; Length 1611;
Best Local Similarity 72.2%; Pred. No. 29;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
OY 2 gctcagsgaatcgaagtcagactttgacg 37
Db 998 GCCTAATATATATATCCGATGGAAGATGTTTGGCG 963
```

RESULT 5
AAAX20344/c
ID AAAX20344 standard; DNA; 976 BP.
XX
XX AAAX20344;
DE 04-MAY-1999 (first entry)

DE Borrelia burgdorferi polynucleotide sequence #97.

KW Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
KW endemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
KW infection; diagnosis; characterisation; detection; ds.

XX Borrelia burgdorferi.

PN W09858943-A1.

PD 30-DEC-1998.

PF 18-JUN-1998; 98MO-US12764.

XX 03-SEP-1997; 97US-0057483.

PR 20-JUN-1997; 97US-0050359.

PR 22-JUL-1997; 97US-0053344.

PR 22-JUL-1997; 97US-0053377.

PA (HUMA-) HUMAN GENOME SCI INC.

XX (MEDI-) MEDIMUNE INC.

PI Clayton R, Dougherty BA, Fraser C, Iathigra R, Smith HO;

PI White OR;

XX WPI; 1999-081217/07.

PS Claim 1; Page 1077-1078; 1128pp; English.

XX AAAX20248 to AAAX20402 represent polynucleotide sequences isolated from
CC Borrelia burgdorferi (Bb). Products derived from Bb can be used for
CC the detection, diagnosis, characterisation, prevention and therapy of
CC Bb infections, e.g. Lyme disease. They can also be used for the
CC production of biosynthetic products, e.g. enzymes. Borrelia belongs
CC to a family of motile, spiral-shaped bacteria called Spirochetes.
CC Spirochetes are pathogenic in humans and Borrelia causes epidemic and
CC endemic relapsing fever, and Lyme borreliosis, more commonly known as
XX Lyme disease.

XX Sequence 976 BP; 308 A; 171 C; 94 G; 400 T; 3 other;

Query Match 52.1%; Score 19.8; DB 20; Length 976;
Best Local Similarity 77.4%; Pred. No. 33;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

CC morphogenesis, recombination, metabolic or catabolic pathway
CC engineering. Using ESTs provides several advantages over genomic or
CC random cDNA clones including elimination of redundancy as one spot on an
CC array equals one gene or open reading frame, and organisation of the
CC microarrays based on function of the gene products to facilitate
CC analysis of the results. AAF07478 to AAF11247 represents ESTs from
CC *Fusarium venenatum*; AAF11248 to AAF11853 represents ESTs from
CC *niger*; AAF11854 to AAF14878 represents ESTs from *Aspergillus oryzae*; and
CC AAF14879 to AAF15337 represents ESTs from *Trichoderma reesei*, which are
CC all specifically claimed in the present invention.
XX
SQ Sequence 729 BP; 173 A; 194 C; 187 G; 175 T; 0 other;

Query Match 53.7%; Score 20.4; DB 21; Length 729;
Best Local Similarity 80.0%; Pred. No. 17;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 6 taagagaatcgaagtacgatactttg 35
||||||| ||||| ||| ||
DB 653 TAGAGGAATTGAGTGCACGAACACTCCGCG 624

RESULT 4
AAC42471/c
ID AAC42471 standard; DNA; 1611 BP.
XX
AC AAC42471;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 35686.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.

XX Nucleic acids encoding internal ribosome entry sequences useful for
PT directing protein expression in gene therapy procedures -
XX Disclosure; Page 32-33; 57pp; English.
XX The PITSURE protein kinase gene can be translated to give two isoforms,
CC p110 and p58. Transcription of p58 mRNA was found to start at an
CC internal ribosome entry site (IRES). The IRES element was found to be
CC cell cycle regulated, p58 is produced during the G2/M stage of the cell
CC cycle. The present sequence is the human PITSURE protein kinase isoform
CC alpha2.2. This sequence contains the IRES. The IRES sequence and
CC fragments of it may be used to induce cell cycle dependent initiation
CC of translation in eukaryotic cells. Vectors containing the IRES may be
CC used for the preparation of compositions for the treatment of and/or
CC prevention of a disease by gene therapy. Such diseases may be cancer
CC and restenosis.
XX
SQ Sequence 2471 BP; 697 A; 601 C; 809 G; 364 T; 0 other;

Query Match 57.9%; Score 22; DB 21; Length 2471;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 17 gaagtcagcattcttgcgc 38
Db 7 gaagtcagcattcttgcgc 28

RESULT 2
AAAX0280
ID AAAX0280 standard; DNA; 3653 BP.
AC AAAX0280;
XX
XX 04-MAY-1999 (first entry)
XX
XX
DE Borrelia burgdorferi polynucleotide sequence #33.
XX
XX
KW Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
KW infection; diagnosis; characterisation; detection; ds.
XX
OS Borrelia burgdorferi.
XX
PN WO9858943-A1.
XX
PD 30-DEC-1998.
XX
PF 18-JUN-1998; 96WO-US12764.
XX
XX
PR 03-SEP-1997; 97US-0057483.
PR 20-JUN-1997; 97US-0050359.
PR 22-JUL-1997; 97US-0053344.
PR 22-JUL-1997; 97US-0053377.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
PA (MEDI-) MEDIMUNE INC.
XX
XX Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
PI White OR;
XX
XX WPI: 1999-081217/07.
XX
XX
PT New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the detection, diagnosis, characterisation, prevention
PT and therapy of infections, particularly Lyme disease
XX
XX Claim 1: Page 994-996; 1128pp; English.
XX
XX AAAX0280 to AAAX0402 represent polynucleotide sequences isolated from
XX Borrelia burgdorferi (Bb). Products derived from Bb can be used for

CC the detection, diagnosis, characterisation, prevention and therapy of
CC Bb infections, e.g. Lyme disease. They can also be used for the
CC production of biosynthetic products, e.g. enzymes. Borrelia belongs
CC to a family of motile, spiral-shaped bacteria called Spirochetes.
CC Spirochetes are pathogenic in humans and Borrelia causes epidemic and
CC endemic relapsing fever, and Lyme borreliosis, more commonly known as
CC Lyme disease.
XX
SQ Sequence 3653 BP; 1587 A; 348 C; 690 G; 1027 T; 1 other;

Query Match 54.7%; Score 20.8; DB 20; Length 3653;
Best Local Similarity 78.1%; Pred. No. 15;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
OY 3 cctagaggaattcgaatgcattcttg 34
Db 3220 ccttagagtaacaattagtgagatacttgg 3251

RESULT 3
AAFI1951/C
ID AAFI1951 standard; CDNA; 729 BP.
XX
XX
AC AAFI1951;
XX
XX 13-MAR-2001 (first entry)
XX
XX
DE Aspergillus oryzae EST SEQ ID NO:4474.
XX
XX
KW Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS Aspergillus oryzae.
XX
XX WO200056762-A2.
XX
XX 28-SEP-2000.
XX
XX 22-MAR-2000; 2000WO-US07781.
XX
XX 22-MAR-1999; 99US-0273623.
XX
XX (NOVO) NOVO NORDISK BIOTECH INC.
XX (NOVO) NOVO NORDISK AS.
XX
PI Berka RM, Rey MW, Shuster JR, Kaupinen S, Clausen IG, Olsen PB;
XX
XX WPI: 2000-594572/56.
XX
XX
PT Monitoring differential expression of genes in filamentous fungal cells
PT uses fluorescence-labeled nucleic acids isolated from the cells and a
PT substrate of expressed sequence tags -
XX
XX
XX Claim 88; Page 1910; 3161pp; English.

CC The present invention describes a method for monitoring differential
CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring
CC the global expression of genes from FF cells allows the production
CC potential of the microorganisms to be improved. New genes may be
CC discovered, possible functions of unknown open reading frames can be
CC identified and gene copy number variation and stability can be
CC monitored. The expression of genes can be used to study how FF cells
CC adapt to changes in culture conditions, environmental stress, spore

DR WPT; 2000-499331/44

Boletia burgdorfii
 Human gene signature
 133/Morine IFN-gam
 Vaccinia J31 promom
 Nucleotide sequenc
 Arabidopsis thaliana
 Human polynucleoti
 Nucleotide sequenc
 Plasmodium yoelli
 Drosophila melanog
 DNA encoding OMP-1
 H. pylori HPS084
 Helicobacter polytyp
 H. pylori CMO 765
 445 bp fragment of
 Human CDNA clone
 Human lung tumor
 Human CDNA clone
 Peroxidase related
 Arabidopsis thaliana
 Peroxidase gene.
 PER peroxidase re
 Arabidopsis thaliana
 Human CDNA sequenc
 Nucleic acid encod
 Human variant Pcep
 Human high affinity
 Human beta subunit
 Human high affinity
 Human adenosine re
 Human IGSRB gene
 Human FCER1 beta c
 Human high affinity
 Human adenosine re

VERSION		U04816.1	GI:507159	
KEYWORDS				
SOURCE		human.		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS		1 (bases 1 to 2471)		
TITLE		Xiang,J., Lathi,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.		
JOURNAL		Molecular cloning and expression of alternatively spliced PITSURE protein kinase isoforms		
MEDLINE		J. Biol. Chem. 269, 15786-15794 (1994)		
REFERENCE		94253170		
AUTHORS		2 (bases 1 to 2471)		
TITLE		Kidd,V.J.		
JOURNAL		Direct Submission		
MEDLINE		Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St., Memphis, TN 38101, USA		
FEATURES		Location/Qualifiers		
SOURCE		1..2471		
		/organism="Homo sapiens"		
		/db_xref="taxon:9606"		
		/map="fp36-2"		
		/cell_line="HeLa S3"		
		/cell_type="epitheloid"		
		/tissue_type="cervix"		
		/clone_lib="HeLa cDNA library"		
		/dev_stage="adult"		
mrna		1..2471		
CDS		112..2445		
		/codon_start=1		
		/product="PITSURE alpha 2-2"		
		/protein_id="AAI19582.1"		
		/db_xref="GI:507160"		
		/translation="MGDEKSWKVITLDELIOEKRRRKEQEAELRLKNSDDRDSKRDLSEBELRDHCMEITIRNPPYREDSDMEDGSEDLSIAIKPQOAMRKRYHHKKLEERKEKHARREKERREKRRRHREDOOKARRHEMOKREMAREHSRRRGDYGVALFRRLQLAKERRERREKMRQOEQRDEQKRRERAERREKRREREVSAAHTMPREDSOKVASHMSRSPPRPREFELGGARKVKPEKKEBERLDSLDDISERTISSAESSASGSSEEEEBEESBEGSTSESEEEEEEEETGSNSEAEQSAAEEVESEEMSEDEERENENHLIVYESRFDRDSGESEAEVEEGTPOSSALTEGDVPDPSPALPIELKOELPYLPALOGCSVEEDPOCLRIIEEGYGVVYRAKDKTDEIVALKRLMKREKKEGPRTISTIREINTLIKOHPIVIYREIYVGNMDKIYIMVNYEHKLKSLMETWKPPLPGEVVTLMIQILRGYKHLHDNLILRDLTSLULSHAGILKVGPGFLAKEYSPKATYPVVVTLMYRAPPELLLAGKEYSTAVDMWSVCIFGELLTOPLPFGKSKEIDINKFVLGTGPSEKIMPYSGSELPAVKMTSRHPYNLRKFGLALSDQGFDLMLNKFTLYPGRRISAEDGKHBYFPRETPLPIDSMFPTMPAPASEQQRVKRGTSPPRPBEGGLGYSOLGDDDLKETGPHLTITNGASASAAGCFSLKF"		
BASE COUNT		697 a 601 c 809 g 364 t		
ORIGIN				
Query Match		57.9%; Score 22; DB 9; Length 2471;		
Best Local Similarity		100.0%; Pred. No. 1.1e+02;		
Matches		22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	17	gaagtacgatacttttgagcg	38	
Dd	7	GAAAGTAGCATACTTTTGCGGC	28	
RESULT 14				
HSU07704				
LOCUS		HSU07704 2471 bp mrna PRI 08-JUL-1994		
DEFINITION		Human protein kinase PITSURE isoform PBETA21 (P58DCDZL1) mRNA,		
ACCESSION		complete cds.		
VERSION		U07704		
KEYWORDS		U07704.1 GI:507426		
SOURCE		human.		
ORGANISM		Homo sapiens		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		


```

gene
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="1"
  /map="1p36.3"
  1..2349
  /gene="CDC21.1"
  /gene="CDC21.1"
  /gene="CDC21.1"
  /note="corresponding genomic sequence deposited as GenBank
  Accession Numbers AF080678-Af080688, AF092429, and
  AF092430"
  /codon_start=1
  /product="PITSLRE protein kinase alpha SV11 isoform"
  /protein_id="AAC72082.1"
  /db_xref="GI:3850314"
  /translation="MREYSDKVKASHMSRSPRPRLFGDGRKPKREKMEERD
  LLSDDISDRKTSNAESSASGSGSEEEENEHLVPESTRDRSGSEAEAE
  BEETGNSSEASAEVSESESESESESESESESESESESESESESESESE
  VGGTPQSSALTEGDYVPSALSPTELKQPLPKPLQGRVSEFQCLNRIEETG
  IGAVYRAKDKTEIDVALKRLKMEKEGFPITSLREINTILKAQHPNIVYREIVG
  SNMDKITYVMNVYEHDLKSLMETMKQPLPGEVKTIIMTLRGVHLHDNVLHRDLK
  TSNLLSHAGILKVGDFGLAREYSPKLAATPYVVTIMYRAPELLGAKKESTAVDMK
  SYGCTGEGFLTKQPLPGKSEIDQNKVFKPDGSEKIMWGYSELPAVKKMTISEHP
  YNNLKRKRGALSDGEPDLANKFTYFPGRIASADGLKHEFTREPLIDPSMFTW
  PAKSEQVYVKRSTSPREPGEGLGYSQLDDDLKETGFLHTTNOGASAGPGFSLKP"

BASE COUNT      668 a      577 c      748 g      356 t

ORIGIN
Query Match      57.9%; Score 22; DB 9; Length 2349;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 17 gaatgacgatactttggcgc 38
|||||
Db 7 GAAGTGACGATCTTTGGCGC 28

RESULT 10
LOCUS      HSU04817      2444 bp      mRNA      PRI      08-JUL-1994
DEFINITION Human protein kinase PITSLRE alpha 2-3 mRNA, complete cds.
ACCESSION  U04817
VERSION     U04817.1 GI:507161
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 2444)
AUTHORS     Xiang,J., Lahti,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.
TITLE       Molecular cloning and expression of alternatively spliced PITSLRE
protein kinase isoforms
JOURNAL     J. Biol. Chem. 269, 15786-15794 (1994)
MEDLINE     94253170
REFERENCE   2 (bases 1 to 2444)
AUTHORS     Kidd,V.J.
TITLE       Direct Submission
JOURNAL     Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's
Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St.,
Memphis, TN 38101, USA

FEATURES
source
  1..2444
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /map="1p36-2"
  /cell_line="Hela S3"
  /cell_type="epitheloid"
  /tissue_type="cervix"
  /clone_lib="Hela cDNA library"
  /dev_stage="adult"
  1..2444
  112..2418
  mRNA
  CDS

```

```

/codon_start=1
/product="PITSLRE alpha 2-3"
/protein_id="AAA19583.1"
/db_xref="GI:507162"
/translation="MDEKDSWKVKTLDELLOBKRRKBOEKAEILRLKNSDDRDK
RDLSEGLDHCMEITIRNSPYRREDMSDRDEEDSLAIKPQOMSREKVHNRKD
EKREKKHARVKKRREHRRKRREODKAREWRORREMAFRHSRENDLEOL
EKREKKRKEOOKBOEOKERERAEERKREARREVSAAHRTMREYSDKVAS
HMSRSPRPRLFGDGRKPKREKMEERDLSLODISDRKTSNAESSASG
SGSESESESESESESESESESESESESESESESESESESESESESESE
ERENENHLVVPESRFDROSGESESEAEVEEGTQSSALTEGDYVPSALPTELK
QELPKYLPALQGRVSEFQCLNRIEEGYGVYRAKDKTEIDVALKRLKMEKEG
FPITSLREINTILKAQHPNIVYREIVGSNMDKITYVMNVYEHDLKSLMETMKQPL
PGEVKTIIMTLRGVHLHDNVLHRDLKTSNLLSHAGILKVGDFGLAREYSPKLA
ATPYVVTIMYRAPELLGAKKESTAVDMKSYGCTGEGFLTKQPLPGKSEIDQNKV
KDIQTPSEKIMWGYSELPAVKKMTISEHPYNNLKRKRGALSDGEPDLANKFTYFPG
RISAEGLKHEFTREPLIDPSMFTWPAKSEQVYVKRSTSPREPGEGLGYSQLDDDLKETG
FLHTTNOGASAGPGFSLKP"

BASE COUNT      693 a      596 c      798 g      357 t

ORIGIN
Query Match      57.9%; Score 22; DB 9; Length 2444;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 17 gaatgacgatactttggcgc 38
|||||
Db 7 GAAGTGACGATCTTTGGCGC 28

RESULT 11
LOCUS      HSU07705      2465 bp      mRNA      PRI      08-JUL-1994
DEFINITION Human protein kinase PITSLRE isoform PBETA22 (p58CDC21.1) mRNA,
complete cds.
ACCESSION  U07705
VERSION     U07705.1 GI:507428
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 2465)
AUTHORS     Xiang,J., Lahti,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.
TITLE       Molecular cloning and expression of alternatively spliced PITSLRE
protein kinase isoforms
JOURNAL     J. Biol. Chem. 269, 15786-15794 (1994)
MEDLINE     94253170
REFERENCE   2 (bases 1 to 2465)
AUTHORS     Kidd,V.J.
TITLE       Direct Submission
JOURNAL     Submitted (14-MAR-1994) Vincent J. Kidd, St. Jude Children's
Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St.,
Memphis, TN 38101, USA

FEATURES
source
  1..2465
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="1"
  /map="1p36-2"
  /cell_line="Hela S3"
  /cell_type="epitheloid"
  /tissue_type="cervix"
  /clone_lib="Hela cDNA library"
  /dev_stage="adult"
  1..2465
  112..2439
  /gene="p58CDC21.1"
  112..2439
  /gene="p58CDC21.1"
  /function="protein kinase"
  /codon_start=1

```

```

SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 2226)
AUTHORS     Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,
            Ambros,P.F. and Kidd,V.J.
TITLE       Duplication of a genomic region containing the Cdc2L1-2 and
            MPM21-22 genes on human chromosome 1p36.3 and their linkage to D122
            Genome Res. 8 (9), 929-939 (1998)
JOURNAL     98424414
MEDLINE     2 (bases 1 to 2226)
REFERENCE   Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,
AUTHORS     Ambros,P. and Kidd,V.J.
TITLE       Direct Submission
JOURNAL     Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's
            Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
FEATURES    Location/Qualifiers
            1..2226
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="1"
            /map="1p36.3"
            1..2226
            /gene="CDC2L1"
            620..2200
            /note="corresponding genomic sequence deposited as GenBank
            Accession Numbers AF080678-AF080688, AF092429, and
            AF092430"
            /product="PITSLRE protein kinase alpha SV4 isoform"
            /protein_id="AAC72078.1"
            /db_xref="GI:3850306"
            /translation="MEERDLISLDQDISSEKRTSSAESGSAAGSGSEEESEEE
            EEEGSGTSESESEEESEEEETGSGSEASQSAEEVSEESSEDEEENENHLLV
            VPESRFDSDGSESEEAEEVEEGTQPSALTEDYVPSFALSPIELKQELPYLFL
            QGCRSVEFQCLNRIEETGYVYRAKDKTDEIVALKRLKMEKEGEPITSLREIN
            TILKAOHPNIVTVREIYVSGNMKIYVMVYVHDLKSLMETKOPDLGEEVTLMO
            LILGKVLHNDNMLHLDIKTNSLISAGLIKVDGCIAGTSGSPKATAPYVVTIAY
            RABELLGAKETSTAVDMMSVGCIFGELTQKPLFPEKSTIDINKYFKDLGDSKRI
            WPGSELPAVKMTFSEHPYNNLKRREGALLSDQGDLMKFLTYFPGRRIASDEGLK
            HEYRETPLEIDSMPTWPAKSEQGVKRGTSRPEEGIGYSQGLDDDLKKTGPHL
            TTTNCGASAGPFSFKT"

BASE COUNT      638 a      541 c      705 g      342 t
ORIGIN

Query Match      57.9%; Score 22; DB 9; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gaagtacgatactttggcgc 38
|||||
Db 7 GAAGTGACGATACCTTTGGCGC 28

RESULT 8
LOCUS      HSU04818      2340 bp      mRNA      PRI      08-JUL-1994
DEFINITION Human protein kinase PITSLRE alpha 2-4 mRNA, complete cds.
ACCESSION  U04818
VERSION     U04818.1 GI:507163
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 2340)
AUTHORS     Xiang,J., Lahli,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J.
TITLE       Molecular cloning and expression of alternatively spliced PITSLRE
            protein kinase isoforms
JOURNAL     J. Biol. Chem. 269, 15786-15794 (1994)

```

```

MEDLINE     94253170
REFERENCE   2 (bases 1 to 2340)
AUTHORS     Kidd,V.J.
TITLE       Direct Submission
JOURNAL     Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's
            Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St.,
            Memphis, TN 38101, USA
FEATURES    Location/Qualifiers
            1..2340
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /map="1p36-2"
            /cell_line="Hela S3"
            /cell_type="epitheloid"
            /tissue_type="cervix"
            /clone_lib="Hela cDNA library"
            /dev_stage="adult"
            1..2340
            /codon_start=1
            626..2314
            /product="PITSLRE alpha 2-4"
            /protein_id="AAA19584.1"
            /db_xref="GI:507164"
            /translation="MREDYSDKVKASHMSRPPRRREFLSDGRKPKVEKMEERD
            LLSLDQDISSEKRTSSAESGSAAGSGSEEESEEESEEGSTSESESEEESEEE
            EETGNSSEASEQSAEEVSEESSEDEEENENHLLVPSRFDSDGSESEEAEEVGE
            GPSSALTEDYVPSFALSPIELKQELPYLALQGRSVEEERQCLNRIEETGYV
            YRAKDKTDEIVALKRLKMEKEGEPITSLREINTILKAOHPNIVTVREIYVSGNM
            DITVVMNVHDLKSLMETKOPFLPEEVVTLMOQLRGVHLHNDNMLHLDIKTSN
            LILSHAGILKVDLGLAREYSGPLAAVYVVTIAYRAPELLGAKETSTAVDMMSV
            CIEGELTQKPLFPEKSTIDINKYFKDLGDSKRIWPGSELPAVKMTFSRHPYNN
            LKRRFGALLSDQGDLMKFLTYFPGRRIASDEGLKHEYRETPLEIDSMPTWPAK
            SEQGVKRGTSRPEEGIGYSQGLDDDLKKTGPHLTTNCGASAGPFSFKT"

BASE COUNT      666 a      576 c      741 g      357 t
ORIGIN

Query Match      57.9%; Score 22; DB 9; Length 2340;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 gaagtacgatactttggcgc 38
|||||
Db 7 GAAGTGACGATACCTTTGGCGC 28

RESULT 9
LOCUS      AF067517      2349 bp      mRNA      PRI      07-NOV-1998
DEFINITION Homo sapiens PITSLRE protein kinase alpha SV11 isoform (CDC2L1)
ACCESSION  AF067517
VERSION     AF067517.1 GI:3650313
KEYWORDS
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 2349)
AUTHORS     Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,
            Ambros,P.F. and Kidd,V.J.
TITLE       Duplication of a genomic region containing the Cdc2L1-2 and
            MPM21-22 genes on human chromosome 1p36.3 and their linkage to D122
            Genome Res. 8 (9), 929-939 (1998)
JOURNAL     98424414
MEDLINE     2 (bases 1 to 2349)
REFERENCE   Gururajan,R., Lahli,J.M., Grenet,J., Easton,J., Gruber,I.,
AUTHORS     Ambros,P. and Kidd,V.J.
TITLE       Direct Submission
JOURNAL     Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's
            Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
FEATURES    Location/Qualifiers
            1..2349

```

```

repeat_region /rpt_family="Alu"
6900. 6913
/rpt_family="T-rich"
repeat_region 6914. 7198
/rpt_family="Alu"
misc_feature 6928. 6943
/ncore="similar to EST AA760985 (NID:g2809915) nx32907.sl
repeat_region 9140. 9425
/rpt_family="Alu"
repeat_region 9426. 9466
/rpt_family="(TAAA)n"
repeat_region 9859. 9991
/rpt_family="L1"
repeat_region 9992. 10519
/rpt_family="MER1_type"
repeat_region 10520. 11545
/rpt_family="L1"
STS /db_xref="GI:1222793"
complement(12187..12419)
repeat_region 13007. 13164
/rpt_family="MIR"
repeat_region 13312. 13338
/rpt_family="AT_rich"
repeat_region 13905. 13937
/rpt_family="(TTA)n"
repeat_region 13938. 14185
/rpt_family="Alu"
repeat_region 14776. 14800
/rpt_family="AT_rich"
repeat_region 15022. 15327
/rpt_family="Alu"
repeat_region 15551. 15617
/rpt_family="Alu"
misc_feature 16049. 16068
/ncore="match to EST N71502 (NID:g1228214) yw32a12.r1"
repeat_region 17185. 17211
/rpt_family="(TTTTC)n"
repeat_region 17212. 17495
/rpt_family="Alu"
repeat_region 18024. 18155
/rpt_family="Alu"
repeat_region 18190. 18236
/rpt_family="L2"
repeat_region 18277. 18493
/rpt_family="L2"
repeat_region 18942. 19025
/rpt_family="MIR"
repeat_region 19608. 19920
/rpt_family="Alu"
repeat_region 20351. 20475
/rpt_family="Alu"
misc_feature 20586. 20820
/ncore="match to EST AA326909 (NID:g1979280)"
repeat_region 21520. 21689
/rpt_family="L2"
repeat_region 21690. 21809
/rpt_family="Alu"
repeat_region 21810. 22923
/rpt_family="L2"
repeat_region 23090. 23229
/rpt_family="MIR"
repeat_region 23582. 23680
/rpt_family="L1"
repeat_region 23681. 23854
/rpt_family="Alu"
repeat_region 23855. 24135
/rpt_family="Alu"
repeat_region 24136. 24278
/rpt_family="Alu"
repeat_region 24279. 24373
/rpt_family="L1"
repeat_region 24374. 24688
/rpt_family="Alu"

```

repeat_region	24689..24688	/rpt_family="L1"
repeat_region	24889..25286	/rpt_family="L1"
repeat_region	25287..25867	/rpt_family="L1"
repeat_region	25819..25967	/rpt_family="L1"
Query Match	58.4%;	Score 22.2; DB 9; Length 219436;
Best Local Similarity	77.1%;	Pred. No. 87;
Matches	27; Conservative	0; Mismatches 8; Indels 0; Gaps 0;
QY	4	tctagaagaattcgaagtcagatactttgagcgc 38
Db	121342	TCYAGAGCAATTCGAAGTACGATACCTTTTCATGC 121376
RESULT	6	
LOCUS	CDCC2L1S01	
DEFINITION	CDCC2L1S01 1060 bp	DNA
ACCESSION	Homo sapiens PITSIRE protein kinase (CDCC2L1) gene, exon 1.	PRI
VERSION	AF080684	
KEYWORDS	AF080684.1	GI:3978425
SEGMENT	1 of 13	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
AUTHORS	1 (bases 1 to 1060)	
TITLE	Gururajan, R., Lahli, J.M., Grenet, J., Easton, J., Gruber, I., Ambros, P.F. and Kidd, V.J.	
JOURNAL	Duplication of a genomic region containing the Cdc2L1-2 and MMP21-22 genes on human chromosome 1p36.3 and their linkage to D12Z2	
REFERENCE	Genome Res. 8 (3), 929-939 (1998)	
MEDLINE	98424414	
AUTHORS	2 (bases 1 to 1060)	
TITLE	Gururajan, R., Lahli, J.L., Grenet, J., Easton, G., Gruber, I., Ambros, P. and Kidd, V.J.	
JOURNAL	Direct Submission	
REFERENCE	Submitted (27-JUL-1998) Tumor Cell Biology, St Jude Children's Research Hospital, 332 N. Lauderdale St, Memphis, TN 38105, USA	
FEATURES	Location/Qualifiers	
source	1..1060	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="1"	
	/map="1p36.3"	
	<398..496	
exon	/gene="CDCC2L1"	
	/number=1	
BASE COUNT	205 a	299 c 341 g 215 t
ORIGIN		
Query Match	57.9%;	Score 22; DB 9; Length 1060;
Best Local Similarity	100.0%;	Pred. No. 1.1e+02;
Matches	22; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	17	gaagtcagatactttgagcgc 38
Db	404	GAGTGACGATCTTTGGCGC 425
RESULT	7	
LOCUS	AF067513	
DEFINITION	AF067513 2226 bp	mRNA
ACCESSION	Homo sapiens PITSIRE protein kinase alpha SV4 isoform (CDCC2L1)	PRI
VERSION	AF067513	
KEYWORDS	AF067513.1	GI:3850305

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

FEATURES
Source Location/Qualifiers

1. 41984
/organism="Caenorhabditis briggsae"
/strain="GuJARat G16"
/db_xref="taxon:6238"
/clone="G42E09"
ORIGIN

BASE COUNT 13435 a 7657 c 8328 g 12564 t

Query Match 60.0%; Score 22.8; DB 3; Length 41984;
Best Local Similarity 79.4%; Pred. No. 50;
Matches 27; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1 tgcctagaggaatcgagctagactcttgg 34
||||| ||||| ||||| ||||| |||||
Db 3820 tgcctagaggaatcgagctagactcttgg 3853

RESULT 5
LOCUS AC005046 219436 bp DNA PRI 02-OCT-2000
DEFINITION Homo sapiens BAC clone CTB-13F3 from 7q22, complete sequence.
ACCESSION AC005046
VERSION AC005046.3 GI:6094632
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 219436)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
REFERENCE 2 (bases 1 to 219436)
AUTHORS Leonard, S. and Miller, N.
TITLE The sequence of Homo sapiens BAC clone CTB-13F3
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 219436)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (12-JUN-1998) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108 USA

REFERENCE 4 (bases 1 to 219436)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (29-JUL-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 5 (bases 1 to 219436)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 6 (bases 1 to 219436)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (02-OCT-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Oct 22, 1999, this sequence version replaced gi:5649376.

Center: Washington University Genome Sequencing Center
Center code: WUSC
Web site: <http://genome.wustl.edu/gsc>

Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_RG013F03

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHRT>, send malto:egreen@nhgri.nih.gov, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

Clone CTB-13F3 is from the first release of the human BAC library CTB-978SK-B. The library contains cloned DNA from the male fibroblast cell line 978SK. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). This clone is available from Research Genetics, Inc. (<http://www.resgen.com>).

VECTOR: pBel0BAC11

Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is CTA-364P16, 200 bp overlap. Actual start of this clone is at base position 83442 of CTA-364P16; actual end is at base position 219436 of CTB-13F3.

The clone CTB-13F3 contains three transposons which have been omitted from the submitted sequence. The transposons should insert after base position 14419, 66205, and 79419 respectively.

FEATURES

Source

1. 219436
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q22"
/clone="CTB-13F3"
/clone_lib="CTB-978SK-B"
2179. .2254
/rpt_family="MIR"
2268. .2541
/rpt_family="Alu"
2948. .3174
/rpt_family="MIR"
3547. .3583
/rpt_family="AT_rich"
3705. .3730
/rpt_family="CAAAA)n"
4127. .4151
/rpt_family="AT_rich"
4154. .4271
/rpt_family="Alu"
4793. .4900
/rpt_family="MER1_type"
6540. .6582
/rpt_family="T-rich"
6583. .6899
repeat_region

Summary Statistics	
Contig length:	184447
Phrap values in estimate:	183714
Average error rate (BCR-Phrap estimate):	0.000257955
Fraction of Phrap values less than 40 :	0.0453001
Number of consensus changing edits:	30
Number of N's in consensus :	0

Position	Original+Context	Consensus changing edits	Edited+Context
7134	gatatatgcg(n)taagtttttt	gatatatgcg(t)taagtttttt	gatatatgcg(t)taagtttttt
18131	cccaatagctg(n)caacaacac	cccaatagctg(a)caacaacac	cccaatagctg(a)caacaacac
25666	lactaaact(n)caaaaaaaa	lactaaact(t)caaaaaaaa	lactaaact(t)caaaaaaaa
34432	actggtatcc(n)ccatbacacg	actggtatcc(a)ccatbacacg	actggtatcc(a)ccatbacacg
44661	gcccaaacac(n)ggggagacag	gcccaaacac(t)ggggagacag	gcccaaacac(t)ggggagacag
44662	gcccaagacn(n)ggggagacg	gcccaagacn(g)ggggagacg	gcccaagacn(g)ggggagacg
93106	ttaaaaagc(n)aacacgcta	ttaaaaagc(a)aacacgcta	ttaaaaagc(a)aacacgcta
93108	ggaacacagc(n)taagacaga	ggaacacagc(a)taagacaga	ggaacacagc(a)taagacaga
93155	ttgaaaagc(n)caattcaaa	ttgaaaagc(a)caattcaaa	ttgaaaagc(a)caattcaaa
133602	tttttttt(g)taacttttt	tttttttt(t)taacttttt	tttttttt(t)taacttttt
133613	acttttttt(n)tttaacttt	acttttttt(t)tttaacttt	acttttttt(t)tttaacttt
139615	ggtygtggtg(n)tcacacgt	ggtygtggtg(c)tcacacgt	ggtygtggtg(c)tcacacgt
139646	gcaaaactcc(n)ctnaaaaaa	gcaaaactcc(a)ctnaaaaaa	gcaaaactcc(a)ctnaaaaaa
139949	aaactcgcct(n)aaaaaaaaa	aaactcgcct(a)aaaaaaaaa	aaactcgcct(a)aaaaaaaaa
146974	gtaatgtgac(n)ntgtacata	gtaatgtgac(a)ntgtacata	gtaatgtgac(a)ntgtacata
146975	taagtgtgac(n)tgtagaac	taagtgtgac(a)tgtagaac	taagtgtgac(a)tgtagaac
149511	taagttgaag(n)tgtagaac	taagttgaag(t)tgtagaac	taagttgaag(t)tgtagaac
151700	ctctcgagac(n)cttaagat	ctctcgagac(a)cttaagat	ctctcgagac(a)cttaagat
154316	lataacctta(n)nnctgctct	lataacctta(c)nnctgctct	lataacctta(c)nnctgctct
154318	alacacctta(n)nnctgctct	alacacctta(c)nnctgctct	alacacctta(c)nnctgctct
154319	taaccttaan(n)ncgtgctct	taaccttaan(c)ncgtgctct	taaccttaan(c)ncgtgctct
154419	aaccttaann(n)ctgtctcca	aaccttaann(c)ctgtctcca	aaccttaann(c)ctgtctcca
155088	tttgaatag(t)aaaaaatca	tttgaatag(t)aaaaaatca	tttgaatag(t)aaaaaatca
156538	cggtccactt(n)taggaatac	cggtccactt(t)taggaatac	cggtccactt(t)taggaatac
157070	taaaaaatgc(n)catgcaact	taaaaaatgc(a)catgcaact	taaaaaatgc(a)catgcaact
157110	ttgtcaaaag(n)gcctgtttt	ttgtcaaaag(t)gcctgtttt	ttgtcaaaag(t)gcctgtttt
157188	acaatgttta(n)aaacataag	acaatgttta(c)aaacataag	acaatgttta(c)aaacataag
171922	aaatgaactc(n)gccttttag	aaatgaactc(t)gccttttag	aaatgaactc(t)gccttttag
178161	caatltagcc(n)ggcagtttc	caatltagcc(a)ggcagtttc	caatltagcc(a)ggcagtttc

Distribution of Quality < 40 Bases	
#	Phrap Value Range
10001	* * *
9001	* * *
8001	* * *
7001	* * *
6001	* * *
5001	* * *
4001	* * *
3001	* * *
2001	* * *
1001	* * *
01	* * *

```

FEATURES
source
Version: 1.01  qxf0.
-----
location/Qualifiers
1..184447
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP1-1036f1"
1..68
repeat_region
/rpt_family="MER66B"

```

```
repeat_region
/rpt_family="MIR"
499..524
/rpt_family="(CAAG)n"
repeat_region
/rpt_family="A1MC5"
1160..1585
/rpt_family="1592..1673"
repeat_region
/rpt_family="A1usq"
1676..1725
/rpt_family="(TCA)n"
1727..1760
/rpt_family="AT_rich"
repeat_region
complement(2080..2370)
/rpt_family="A1uv"
complement(2395..2679)
/rpt_family="A1usx"
3700..4548
/rpt_family="L2"
4540..4594
/rpt_family="L2"
4604..4624
/rpt_family="(TGA)n"
5423..5931
/rpt_family="L2"
6112..6466
/rpt_family="M1T1"
complement(6525..6638)
/rpt_family="M1R"
6647..6673
/rpt_family="(TTG)n"
complement(6674..6963)
/rpt_family="A1uvo"
complement(6964..7122)
/rpt_family="M1R"
complement(7138..7440)
/rpt_family="A1uv"
7460..7493
/rpt_family="(A)n"
```

Query Match:	61.6%	Score 23.4;	DB 9;	Length 184447;
Best Local Similarity:	81.8%	Pred. No. 28;		
Matches	27;	Conservative	0;	Mismatches 6; Indels 0; Gaps 0;
Qy	3	ctctagaggaattcgaaatgacgactcttttg	35	
b	74647	ctctagaggaattcgaaatgacgactcttttg	74615	

RESULT	4
CBRG42E09	
LOCUS	CBRG42E09 41984 bp DNA INV 04-NOV-2000
DEFINITION	Caenorhabditis briggsae cosmid G42E09, complete sequence.
ACCESSION	AC084593
VERSION	AC084593.1 GI:11095043
KEYWORDS	HTG.
SOURCE	Caenorhabditis briggsae.
ORGANISM	Caenorhabditis briggsae Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdilitoidae; Rhabdilitidae; Peloderinae; Caenorhabditis. 1 (bases 1 to 41984)
REFERENCE	Washington University Genome Sequencing Center. The C. briggsae Genome Sequencing Project. Unpublished 2 (bases 1 to 41984)
AUTHORS	Waterson, R.
JOURNAL	Direct Submission
REFERENCE	Submitted (04-NOV-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
TITLE	Submitted by:
JOURNAL	Genome Sequencing Center: Department of Genetics, Washington University, St. Louis, MO 63110, USA e-mail: jspliehwatson.wustl.edu
COMMENT	

CC	*	71777	71876:	gap of unknown length
CC	*	71877	84981:	contig of 13105 bp in length
CC	*	84982	85081:	gap of unknown length
CC	*	85082	97556:	contig of 12475 bp in length
CC	*	97557	97656:	gap of unknown length
CC	*	97657	111043:	contig of 13387 bp in length
CC	*	111044	111143:	gap of unknown length
CC	*	111144	122275:	contig of 11132 bp in length
CC	*	122276	122375:	gap of unknown length
CC	*	122376	133263:	contig of 10888 bp in length
CC	*	133264	133363:	gap of unknown length
CC	*	133364	140049:	contig of 6686 bp in length
CC	*	140050	140149:	gap of unknown length
CC	*	140150	147482:	contig of 7333 bp in length
CC	*	147483	147582:	gap of unknown length
CC	*	147583	151952:	contig of 4370 bp in length
CC	*	151953	152052:	gap of unknown length
CC	*	152053	156921:	contig of 4865 bp in length
CC	*	156922	157021:	gap of unknown length
CC	*	157022	158420:	contig of 2399 bp in length
CC	*	158421	159520:	gap of unknown length
CC	*	159521	163612:	contig of 4093 bp in length
CC	*	163613	163712:	gap of unknown length
CC	*	163713	167217:	contig of 3505 bp in length
CC	*	167218	167317:	gap of unknown length
CC	*	167318	169954:	contig of 2637 bp in length
CC	*	169955	170054:	gap of unknown length
CC	*	170055	171664:	contig of 1610 bp in length
CC	*	171665	171764:	gap of unknown length
CC	*	171765	173324:	contig of 1560 bp in length
CC	*	173325	173424:	gap of unknown length
CC	*	173425	175056:	contig of 1633 bp in length
CC	*	175057	175156:	gap of unknown length
CC	*	175157	176359:	contig of 1203 bp in length.
XX				
FH	Key		Location/Qualifiers	
FT	source	1..176359		
FT		/chromosome="3"		
FT		/db_xref="taxon:9606"		
FT		/organism="Homo sapiens"		
FT		/clone="RP11-360G1"		
XX	Sequence 176359 BP; 53121 A; 34541 C; 34264 G; 52495 T; 1938 other.			

Query Match	61.6%	Score 23.4	DB 33	Length 176359
Best Local Similarity	81.8%	Pred. No. 28		
Matches 27	Conservative 0	Mismatches 6	Indels 0	Gaps 0
QY	3	ctctagaggaattcgaattgcagatcattctgg	35	
Db	36576	CTCTAGAGGAATTCAGATCAACATACGTATGG	36544	
RESULT 3				
AC012519/c				
LOCUS	AC012519	184447 bp	DNA	
DEFINITION	Homo sapiens	3 BAC RP11-1036F1 (Roswell Park Cancer Institute Human		
	BAC library)	complete sequence.		
ACCESSION	AC012519			
VERSION	AC012519.9	GI:9558566		
KEYWORDS	HTG.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eumariota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 184447)			
AUTHORS	Muzny,D.M., Adams,C., Bailey,M., Barbara,J., Blankenburg,K.,			
	Bodala,B., Buick,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,			
	Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,			
	David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,			
	Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,			

[illegible]

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: this sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

TITLE Direct Submission
JOURNAL Submitted (10-AUG-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
COMMENT On Aug 10, 2001 this sequence version replaced gi:8699801.
----- Genome Center
Center: University of Washington Genome Center
Web site: <http://www.genome.washington.edu>
Contact: uwgchgs@u.washington.edu
Drafting Center: BCM
----- Project Information
Center project name: chr-3
Center clone name: RP11-360G1 (bc0384)
----- Summary Statistics
Sequencing vector: M13; 108821; 50% of reads
Sequencing vector: plasmid; 108752; 50% of reads
Chemistry: Dye-primer Bodipy; 2% of reads
Chemistry: Dye-terminator ET; 32% of reads
Chemistry: Dye-terminator Big Dye; 66% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 173927 bases at least Q40
Consensus quality: 174208 bases at least Q30
Consensus quality: 174306 bases at least Q20
Insert size: 174350; sum-of-contigs
Quality coverage: 8.0x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 6687: contig of 6687 bp in length
* 6688 6787: gap of unknown length
* 6788 174450: contig of 167663 bp in length.
* Location/Qualifiers
1..174450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="3"
/clone_lib="RP11-360G1"
/clone_lib="RP11-360G1"
/clone_lib="RP11-360G1"
misc_feature 1..6687
/note="assembly_name:Contig18"
misc_feature 6788..174450
/note="assembly_name:Contig19"
BASE COUNT 53199 a 34138 c 34305 g 52708 t 100 others
ORIGIN
Query Match 61.6%; Score 23.4; DB 2; Length 174450;
Best Local Similarity 81.8%; Pred. No. 28;
Matches 27; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
OY 3 cttctaggaatcgaagtcagatcttttg 35
DB 75606 CTTCTAGGAATGAAGTGAACATACGTATG 75574
RESULT 2
AC055709/c
ID AC055709 standard; DNA: HTG; 176359 BP.
XX AC055709;
XX AC055709;
XX AC055709.7
XX
DT 20-APR-2000 (rel. 63; Created)
DT 07-SEP-2000 (rel. 65; Last updated, Version 11)
XX
XX Homo sapiens chromosome 3 clone RP11-360G1, WORKING DRAFT SEQUENCE, 20

DE unordered pieces.
XX HTG; HTGS_DRAFT; HTGS_PHASE1.
KW
XX
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
XX
XX [1]
XX 1-176359
XX
XX Muzny D.M., Adams C., Bailey M., Barbaria J., Blankenburg K., Bodda B.,
XX Bouck J., Bowe S., Brooks A., Buhay C., Bunac C., Burkett C., Burrows J.,
XX Carter M., Chacko J., Chen Z., Cox C., David R., Delgado O., Deshazo D.,
XX Ding Y., Domah-Rashid N., Dugan-Rocha S., Durbin K.J., Fernandez C.,
XX Ferraguto D., Forcun-Tansey J., Frantz P., Ganesh R., Gorrell J.H.,
XX Gorrell L.L., Guevara W., Harris K., Hernandez J., Hodgson A., Hoques M.,
XX Holloway C., Hosack H., Jackson L.E., Jackson L., Jia Y., Jones M.,
XX Kelly S., Kondajewski N., Kong Y., Kovar C., Leal B., Li Z., Lichtarge O.,
XX Liu J., Liu W., Logan O., Lozado R.J., Lu J., Lucier R., Martin R.,
XX Martinez C., McLeod M.P., Mei G., Morgan M., Morris S., Nash S., Nelson A.,
XX Nguyen R., Nguyen N., Nguyen S., Oswald G., Parish B., Paxton S., Payton B.,
XX Perez L., Pu L.L., Quiles M., Reiter D., Rives M., Samuel S., Say J.,
XX Scherer S., Shen H., Simon M., Sparks A., Stamps A., Sugang R.,
XX Tabor P., Taylor T., Vasquez L., Vinson R., Vo Q., Wabash M.,
XX Watling S., Weinstein G., Weinstein I.R., Williamson A., Worley K.,
XX Wren J., Wrensford G., Yu W., Zhou X., Nelson D., Gibbs R.,
XX "Direct Submission";
XX unpublished.
XX
XX [2]
XX 1-176359
XX Worley K.C.;
XX
XX Submitted (18-APR-2000) to the EMBL/GenBank/DBJ databases.
XX Human Genome Sequencing Center, Department of Molecular and Human Genetics,
XX Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
XX
XX
XX On Jun 25, 2000 this sequence version replaced gi:8102160.
XX
XX ----- genome center
XX Center: Baylor College of Medicine
XX Center code: BCM
XX Web site: <http://www.hgsc.bcm.tmc.edu/>
XX Contact: hgsc-help@bcm.tmc.edu
XX ----- Project Information
XX Center project name: HAP0
XX Center clone name: RP11-360G1
XX ----- Summary Statistics
XX Sequencing vector: M13; 108821
XX Chemistry: Dye-primer Bodipy; 4% of reads
XX Chemistry: Dye-terminator Big Dye; 96% of reads
XX Assembly program: Phrap; version 0.990329
XX Consensus quality: 156972 bases at least Q40
XX Consensus quality: 165022 bases at least Q30
XX Consensus quality: 169145 bases at least Q20
XX Estimated insert size: 170012; sum-of-contigs estimation
XX Quality coverage: 0x in Q20 bases; agarose-1P estimation
XX Quality coverage: 3.9x in Q20 bases; sum-of-contigs estimation
XX -----
XX * NOTE: Estimated insert size may differ from sequence length
XX (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
XX * NOTE: This is a 'working draft' sequence. It currently
XX * consists of 20 contigs. The true order of the pieces
XX * is not known and their order in this sequence record is
XX * arbitrary. Gaps between the contigs are represented as
XX * runs of N, but the exact sizes of the gaps are unknown.
XX * This record will be updated with the finished sequence
XX * as soon as it is available and the accession number will
XX * be preserved.
XX 1 31369: contig of 31369 bp in length
XX 31370 31469: gap of unknown length
XX 31470 54720: contig of 23251 bp in length
XX 54721 54820: gap of unknown length
XX 54821 71776: contig of 16956 bp in length

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 4, 2001, 18:54:48 ; Search time 2399.05 Seconds
(without alignments)
261.309 Million cell updates/sec

Title: US-09-915-060-7

Perfect score: 10
Sequence: 1 tgccttagaggaattcgaagtgacgatacttggcgc 38

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 segs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	23.4	61.6	174450	2	AC093115	AC093115 Homo sapi
C 2	23.4	61.6	176359	33	AC055709	AC055709 Homo sapi
C 3	23.4	61.6	184447	9	AC012519	AC012519 Homo sapi
C 4	22.8	60.0	41984	3	CBRG42E09	AC084593 Caenorhab
5	22.2	58.4	219436	6	AC005046	AC005046 Homo sapi
6	22	57.9	1060	9	CDG2L1S01	AF080684 Homo sapi
7	22	57.9	2226	9	AF067513	AF067513 Homo sapi
8	22	57.9	2340	9	HS004818	U04818 Human prote
9	22	57.9	2349	9	AF067517	AF067517 Homo sapi
10	22	57.9	2444	9	HS004817	U04817 Human prote
11	22	57.9	2465	9	HS007705	U07705 Human prote
12	22	57.9	2471	6	AX033422	AX033422 Sequence
13	22	57.9	2471	9	HS004816	U04816 Human prote
14	22	57.9	2471	9	HS007704	U07704 Human prote
15	22	57.9	2477	9	HS004824	U04824 Human prote
16	22	57.9	2486	9	AF067512	AF067512 Homo sapi
17	22	57.9	2500	9	AF067514	AF067514 Homo sapi
18	22	57.9	2525	9	AF067515	AF067515 Homo sapi
19	22	57.9	2533	9	AF067516	AF067516 Homo sapi
C 20	21.8	57.4	18061	9	HSCYTK020	X73501 H. sapiens g
C 21	21.8	57.4	137022	9	AC004231	AC004231 Homo sapi
C 22	21.6	56.8	8171	3	DMVIN	X96601 D.melanogas
C 23	21.6	56.8	17254	3	DROPCXGEN	M74329 Drosophilla
C 24	21.6	56.8	26697	3	DMC103B4	AL009193 Drosophill
C 25	21.6	56.8	36824	8	AP000372	AP000372 Arabidops
C 26	21.6	56.8	38892	2	DMC6209	AL009171 Drosophill
C 27	21.6	56.8	46221	2	AC017684	AL017684 Drosophill
28	21.6	56.8	116524	9	AL138686	AL138686 Human DNA
29	21.6	56.8	175188	2	AL445238	AL445238 Homo sapi
C 30	21.6	56.8	291976	3	AE003423	AE003423 Drosophill
C 31	21.4	56.3	37049	1	MLACBA	Z46257 M.leptae ac
C 32	21.4	56.3	37711	3	CEZK896	Z82288 Caenorhabdi
C 33	21.4	56.3	86825	9	AC006967	AC006967 Homo sapi
34	21.4	56.3	342300	1	MLEPRTN8	AL583924 Mycobacte
C 35	21.2	55.8	652	3	AF176697	AF176697 Circulife
C 36	21.2	55.8	66224	9	AC073884	AC073884 Homo sapi
C 37	21.2	55.8	159691	9	AC025160	AC025160 Homo sapi
C 38	21.2	55.8	172893	2	AC024596	AC024596 Homo sapi
C 39	21	55.3	36783	1	AF078135	AF078135 Leptospi
C 40	21	55.3	46057	9	AC073276	AC073276 Homo sapi
C 41	21	55.3	145603	9	AL157768	AL157768 Human DNA
C 42	21	55.3	169581	2	AC007490	AC007490 Homo sapi
C 43	21	55.3	193539	2	AC022190	AC022190 Homo sapi
C 44	20.8	54.7	2945	3	D83384	D83384 Tetrahymena
C 45	20.8	54.7	5108	1	BB044914	U44914 Borrelia bu

ALIGNMENTS

RESULT 1
AC093115 174450 bp DNA HTG 10-AUG-2001
LOCUS Homo sapiens chromosome 3 clone RP11-360c1, WORKING DRAFT SEQUENCE.
DEFINITION
ACCESSION AC093115
VERSION AC093115.1 GI:15147196
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 174450)
Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J., Ivey,R.G. and
Haugen,E.D.
AUTHORS
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 174450)
Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J., Ivey,R.G. and
Haugen,E.D.

Wed Dec 5 08:45:02 2001

us-09-915-060-1.rge

Page 9

Db 1027 GAGAACTCAGAGGAGGAGCAGAGAGGAGAGAGACCGGACGCAACTCTGAGGAG 1086

Qy 181 gcatcctagagcaatctctgcccgaagaagtaagtgtgaggaagaatg 222

Db 1087 GCATCAGAGCAGTCTGCGCGAAGAGTAACCTGAGGAAGAAATG 1128

Search completed: December 4, 2001, 18:54:30
Job time: 9166 sec

9

	BASE COUNT	693 a	599 c	808 g	365 t	
	ORIGIN					
	Query Match	100.0%; Score 222; DB 9;	Length 2465;			
	Best Local Similarity	100.0%; Pred. No. 1,le-35;				
	Matches 222; Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
Oy	1	gacatcagcgacgcgaaggagaaccacccagtctggcccgaagtccctgttcagacaatacgcgc	60			
Db	901	GACATCAGCGACGACGGAGAAGAACCCAGCTGCCGCAGTCTCTTCAGCAAAATCAGGC	960			
Oy	61	tcaagttcttgaaagaaagagagagaaaagagagagagagagagagagagagagagagagagag	120			
Db	961	TCAAGTTCTTGGAAGAAGAGAGAGAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1020			
Oy	121	gaagaattcagaggaagaaagaaag	180			
Db	1021	GAAGAATTCAAGAGAGAAAG	1080			
Oy	181	gcattcagagcagcttgcgccgaagtaagttagagagaatatg	222			
Db	1081	GCATCAGAGCAGGTCTGCCGAAGAATAAGTAGAGAAAATAG	1122			
	RESULT 14					
	AX033422					
LOCUS	AX033422	2471 bp	DNA	PAT	21-SEP-2000	
DEFINITION	Sequence 3 from Patent WO0044896.					
ACCESSION	AX033422					
VERSION	AX033422.1 GI:10280183					
KEYWORDS	'human.'					
SOURCE	Homo sapiens					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
REFERENCE	1 (bases 1 to 2471) Beyerle,R.,and Cornelis,S. Internal ribosome entry site (IRES), vector containing same and uses thereof Patent: WO 0044896-A 3 03-AUG-2000; VIAMMS INTERNINV INST BIOTECH (BE) ; BEYERLE RUDI (BE) ; CORNELIS STIERID (BE)					
JOURNAL						
FEATURES						
source	Location/Oualifiers					
	1..2471 /organism="Homo sapiens"					
	misc_feature 1..2471 /db_xref="taxon:9606"					
	note="The IRES-activity containing sequence PylstRE protein kinase (PylOplstire) (isoform alfa 2-2)"					
BASE COUNT	697 a	601 c	809 g	364 t		
ORIGIN						
	Query Match	100.0%; Score 222; DB 6;	Length 2471;			
	Best Local Similarity	100.0%; Pred. No. 1.le-35;				
	Matches 222; Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
Oy	1	gacatcagcgacgcgaaggagaaccacccagtctggcccgaagtccctgttcagacaatacgcgc	60			
Db	907	GACATCAGCGACGACGGAGAAGAACCCAGCTGCCGCAGTCTCTTCAGCAAAATCAGGC	966			
Oy	61	tcaagttcttgaaagaaagagagagaaaagagagagagagagagagagagagagagagagagag	120			
Db	967	TCAGTTCTTGGAAGAAGAGAGAGAGAGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1026			
Oy	121	gaagaattcagaggaagaaagaaag	180			
Db	1027	GAAGAATTCAAGAGAGAAAG	1086			
Oy	181	gcattcagagcagcttgcgccgaagtaagttagagagaatatg	222			
Db	1087	GCATCAGAGCAGGTCTGCCGAAGAATAAGTAGAGAAAATAG	1128			

[illegible]

	Db	1000	GAAGAAATCAGACGAGAGGAGCAAGAAGAGGAGAGGAGAGACCCGGACAGCAACTCTGAGGAG	1059
OY	181		gcatacagacgactctgccgaagaagtlaagtggagaagaatg	222
Db	1060		GCATCAGAGCACTGTGCCGAGAGAACTAAGTAGGAAAGATG	1101
RESULT	12			
AF067522				
LOCUS				
DEFINITION	Homo sapiens PITSURE protein kinase beta SV6 isoform (CDC2L2) mRNA,			
ACCESSION	AF067522			
VERSION	AF067522.1			
KEYWORDS	GI:3850323			
SOURCE	.			
ORGANISM	human.			
	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
	1 (bases 1 to 2448)			
TITLE	Guturajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,			
JOURNAL	Ampros,P.F. and Kidd,V.J.			
MEDLINE	Duplication of a genomic region containing the Cdc2L1-2 and			
REFERENCE	MMP21-22 genes on human chromosome 1p36.3 and their linkage to D1Z22			
AUTHORS	Genome Res. 8 (3), 929-939 (1998)			
	2 (bases 1 to 2448)			
TITLE	Guturajan,R., Lahti,J.M., Grenet,J., Easton,J., Gruber,I.,			
JOURNAL	Ampros,P. and Kidd,V.J.			
MEDLINE	Direct Submission			
REFERENCE	Submitted (10-MAY-1998) Tumor Cell Biology, St. Jude Children's			
AUTHORS	Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA			
FEATURES	Location/Qualifiers			
source	1..2448			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="11"			
	/map="1p36.3"			
	1..2448			
gene	/gene="CDC2L2"			
CDS	80..2422			
	/gene="CDC2L2"			
	/note="corresponding genomic sequence deposited as GenBank			
	Accession Numbers AF080689-AF080697 and AF092426-AF092428"			
	/codon_start=1			
	/product="PITSURE protein kinase beta SV6 isoform"			
	/protein_id="AAC72087.1"			
	/db_xref="GI:3850324"			
	/translation="MGDEKDSWKYKTLDTELIOEKRRKEQEAIEIRLKNSDDRSK			
	RDSLSEGLRDHCHEITRNSPYREDSMEDGEEDSLAKPPQMSWKKVHHNRKD			
	KDSEKMKHARVKEREHRERKRREEDOKRAEMEDROKRRMARERSREDRLLEOLE			
	KKRRERKMREQOEOROKREERAERERRAREBARRESVHNRTMEDRYADKYAS			
	WSRPAPRPAPREPFLGDGRKGRCGARPARAPAKPAKLEKMERLLSDIODISPER			
	TSSAEASSAESGSSEEEEBEGBEGSTSESEBEEBEETSGNSEASESQS			
	EEVESEEWSEDEREENENHLVPESRPDRNGSEEAEEVEGGTQQSKALTEGDYV			
	PDSALLPIELKOELPKYLPAIQCSYSEEOCLNRIEEGYGVYRAKDKTEIDYVA			
	LKRILMEKEKGFPTSLREINTLIKOHPIVIYREIVGSSNDKIYIVNVYEHDIA			
	KSLMETMKEOPELVGETIMTOILRGVKLHDMNIILHDLIKTSNLISHGILKVGGE			
	GLADEYSPLKATYPVVYVTQYRARPELLGAKKESTAVDMDSVCIGTELLTQKPLPF			
	GNSELDTINKVFKEGLGPSEKTIWPGYSLELPVVKMTSTSEHPNNURKFFGLSDQG			
	DLNMKPLTVFPGRISAEDGLKHEYFRTPPLPIDPSMPTTPARSEQRKGPSPRSP			
	PEGGIQGYSQLDDDDDKETGFTLTNTNOGASAAGGFSIKF"			
BASE COUNT	694 a	598 c	796 g	360 t
ORIGIN				
Query Match	100.0%; Score 222; DB 9; Length 2448;			
Best Local Similarity	100.0%; Pred. No. 1.le-35;			
Matches	222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
OY	1			
	gacatcacgacagcgagagaaaccagctcgccgagtcctgttcaagcagaatcagc			60
Db	884			
	GACATCAGCGAGCAGGAGGAGAACGACGACTCGCCGAGTCTCTCGTCGACACAATCAGGC			943

ACCESSION	AF067519	
VERSION	AF067519.1	GI:3850317
KEYWORDS		
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (bases 1 to 2439)	
TITLE	Gururajan,R., Tahli,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P.F. and Kidd,V.J.	
JOURNAL	Duplication of a genomic region containing the Cdc2l-2 and MPM21-12 genes on human chromosome 1p36.3 and their linkage to D12Z	
REFERENCE	Genome Res. 8 (9), 929-939 (1998)	
MEDLINE	98424414	
REFERENCE	2 (bases 1 to 2439)	
AUTHORS	Gururajan,R., Tahli,J.M., Grenet,J., Easton,J., Gruber,I., Ambros,P. and Kidd,V.J.	
TITLE	Direct Submission	
JOURNAL	Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's	
FEATURES	Research Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA	
SOURCE	1..2439	Location/Qualifiers

gene
CD5

1. 2439
/gene="CDC2L2"
80. 2413
/gene="CDC2L2"
/note="corresponding genomic sequence deposited as GenBank
Accession Numbers AF080689-AF080697 and AF092426-AF092428"
/codon_start=1
/product="PitSLRE protein kinase beta SV1 isoform"
/protein_id="AAC72084.1"
/db_xref="GI:3850318"
/translation="MGDEKSDMKVLTDEILQEKRRRQDEKAELIKLNSDDRSK
RDSIEEELDHCMELIRNSPYRREDSMDRCGEDSLAIKPKQOAMTEKYHRKD
EKREKCRHHSHSAEGGKHARKVKEHEERKRHRHEDQKRRREKQRRREMRHESR
REKRELEQLERKREKRRRQKQEKQEKQERPRARERKREKREKREKREKREKREK
DYSDVKYASWMSRSSPPRRPRRPELQGRKRPVKEKMEEDLSDLDISDSEKRTTS
AESSSAEGSGSGSEEEEEEEBEGSTSESEEEEEEEEGTNSSEASQSAEIV
SEESMESEDERENENHLVVPESRFDSSESEAEVEEGTPQSSALTEGVDPDS
PALPLELQKQELPKYPLALQGCRSVEFQCLNIEEGTVGVYRAKDKTDEIYALKR
LAKMEKEKEGPITSLREINTILKAOQHNIYVIREIYVGSNMOKYIYMNVVEHDKSL
MEMKQOPLFGEVKTLMIOILRGVKLHDHMTILRHDLKTSNLLISHGILIKVDEGFLA
MEMGSPFLKATYPVYVQWQWRAPELLIGAKKYSTAVDMWMSGCTFGELLTPKPLPFGMS
EIDQINKVEKEITPSEKIMPGYSEL.PYVYKKMTFSEHPINLKKRFKCALLSDQGFIDL
GLITITPPGRIRISAEDELKHYRETFPLDPSMFLPMPYPAKSDQRYVKRTGTSPPRPEG
GLTYSIQGDGDDLEKTEGTHLTYTQGAASAPGFSLP"

BASE COUNT
BRIGIN

695 a 592 c 790 g 362 t

[illegible]

RESULT	11			
LOCUS	HSU04817			
DEFINITION	Human protein kinase PITSURE alpha 2-3 mRNA, complete cds.			
ACCESSION	U04817			
VERSION	U04817.1			
KEYWORDS				
SOURCE				
ORGANISM	human.			
	Homo sapiens			
	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:			
	Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.			
REFERENCE	1 (bases 1 to 2444)			
AUTHORS	Xiang,J., Lantl,J.N., Grenel,J.A., Easton,J.B. and Kidd,V.J.			
TITLE	Molecular cloning and expression of alternatively spliced PITSURE protein kinase isoforms			
	J. Biol. Chem. 269, 15786-15794 (1994)			
JOURNAL	94253170			
MEDLINE	2 (bases 1 to 2444)			
REFERENCE	Kidd,V.J.			
AUTHORS	Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's			
TITLE	Research Hospital, Tumor Cell Biology, 332 N. Lauderdale St.,			
JOURNAL	Memphis, TN 38101, USA			
	Location/Qualifiers			
FEATURES				
source	1..2444			

```

MRNA
112. 2418
CDS
1. .2444
/codon_start=1
/product="PITSLRE alpha 2-3"
/protein_id="AA19585.1"
/db_xref="GI:507162"
/translation="MGSDEKSWKVTLDLDELQEKRRKEDEKAEIKRLKNSDQDSK
EKDLEEGLRDHCMEITIRNSPYRREDSMEDRGSEDSLSIKPQOQSRRKRVHNRKD
ERKSEKHARVKKKEHEHRRKRRRESDQKARRWEMEQRRREARRRERDRLEDQ
ERKNEBERKMRKEQEQDQREKERRAEKREKREARREVSANHRTKMSDDVKAESG
HMSKSPRPREFRLDQGRPKVKEEMKEEDLSDLODTSDESRKTSASSASAEESG
SGSEEEEEEEEEEEEGSTSPSEEEEEEEEEEEEGTSSSEDSQSAEVSSEEMSESD
EREEMENILVYPESGRPDQSGSEAEAEVEEGTSSQSALEGGVSDPSLSTIELK
QELPKLTALQSGSVSEPCQINRIIEGTGVVYRARDKRDELVALIKRLKMEKEKGG
PILTSLEINTILKAOHNPITYVREIYVYVSEIMKDYIYMYNVEDLISLMTKMOPL
PGEVKLTMIQLIRGVKHLHDHMMIHRDLKTSNLSNAGILLKVDGEGLARSTGSLAA
YTPPYVYLVMTIRPGLLIGAKRTSTAYVMWSGCLFGELLYTLPEKSSDLOINKYFA
KLDGTPSEKIMPGISELPAAVKMKTSRHPINNLKRGALISDGGFDLMNKLITYPPG
RRISAEQDLHNEYRFRPLPLDPSMFPYPAKNSQVKKVAKRGTSPRPEGGLGYSLQSD
DLKELEGHLLTITNOGSAAGPGSLF"
BASE COUNT
693 a 596 c 798 g 357 t
ORIGIN

```

[illegible]

2 (bases 1 to 2362)

1. Y L L L L C
 2. P L L L L C

FEATURES		Location/Qualifiers	
source		1..2323	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/chromosome="1"	
		/map="1p36.3"	
gene		1..2323	
		/gene="CDC2L2"	
		1104..2297	
CDS		/gene="CDC2L2"	
		/note="Corresponding genomic sequence deposited as GenBank Accession Numbers AF080689-AF080697 and AF092426-AF092428"	
		/codon_start=1	
		/product="PITSLRE protein kinase beta SV8 isoform"	
		/protein_id="AAC72083.1"	
		/db_xref="GI:3850316"	
		/translation="MKNEKKTKTSMLEPQSHGSEITPGRVKKORKKMGKSGVEEFOCLNRIEKGTCGVYVYRAKDKKTDTEIVALKRLKMKKEKRGPTITSLREINTLTKKQHPVIVREIVYGSNMOKIYIVNIVYEHDLKSLMETKQPLPEEVTLMQLRGVKNLHDNMI LNRDKTSNLLSHAGILKVGDFGLAEVSGSPLKAYIPVVVOTWYRAPELLAKERS TAYDMWVGICIFGELITQKPLFPGNSEIDQINKYFKLGPSEKTPWYSGLPLVYKMM TSEHPYNNLRKRFGLISQGGFDLMKFLTFYFGRRISAEQGLKHEYFRETPLPIDP SMPTTPYAKSEQGRVKKRGTSPPREGSLGYSQGLDDDLKELGFLUTTTNQGASAGPG FSLK"	
BASE COUNT	673 a 550 c 754 g 346 t		
ORIGIN			
Query Match	100.0%; Score 222; DB 9; Length 2323;		
Best Local Similarity	100.0%; Pred. No. 1,1e-35;		
Matches 222; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 gacatcagcgacgacgagagaaagacacgctcgccgagtcctcgtcagcagaatcagc 60		
DB	875 GACATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 934		
QY	61 tcaggtcttcggagc 120		
DB	935 TCAGGTTCTTGAGGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 994		
QY	121 gaagatcagcagcaactctgaggag 180		
DB	995 GAAGATCATCAG 1054		
QY	181 gcatcagagcagctcgtccggaagaagtaagtgtgagagagagagagagagagag 222		
DB	1055 GCATCAGAGCACTCTGCCGAAGAAGTAAGTGAAGAGAAAGTAAT 1096		
RESULT 5			
LOCUS	AF067525 2329 bp mRNA PRI 07-NOV-1998		
DEFINITION	Homo sapiens PITSLRE protein kinase beta SV13 isoform (CDC2L2)		
ACCESSION	AF067525		
VERSION	AF067525.1 GI:3850329		
KEYWORDS			
SOURCE	Human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 2329)		
TITLE	Gururajan, R., Lahli, J.M., Grenet, J., Easton, J., Gruber, I., Ambros, P.F. and Kidd, V.J.		
JOURNAL	Duplication of a genomic region containing the Cdc2L1-2 and MRP21-22 genes on human chromosome 1p36.3 and their linkage to D12Z		
MEDLINE	Genome Res. 8 (9), 929-939 (1998)		
REFERENCE	98424414		
AUTHORS	2 (bases 1 to 2329)		
TITLE	Gururajan, R., Lahli, J.M., Grenet, J., Easton, J., Gruber, I., Ambros, P. and Kidd, V.J.		
JOURNAL	Submitted (18-MAY-1998) Tumor Cell Biology, St. Jude Children's		

FEATURES

source	Hospital, 332 N. Lauderdale St., Memphis, TN 38105, USA
1..2329	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/chromosome="1"	
/map="1p36.3"	
1..2329	
/gene="CDC2L2"	
68..1654	
/note="CDCL2"	
/gene="corresponding genomic sequence deposited as GenBank accession numbers AF080689-AF080697 and AF092426-AF092428"	
/codon_start=1	
/product="PITSLRE protein kinase beta SV13 isoform"	
/protein_id="AAC72090.1"	
/db_xref="GI:3850330"	
/translation="METIRKNSPYRREDSDMEDGDEDSLAIKPPQSKMKETKHNRKDEREKWKRAVKEHEHERKRRIHREODAKREWEROKREMAKESHRSREGNDGFLEFDLDLOLERKERERKMRBOOEOROEOKERRAEERKEREAREARVSANHTMRFLYRDYDOLSHMSRRPAPPREPRLGDGRKPVKEEMERDLSDLOISERTSSAESASGESEEEEEEEDDEEFGSTSESEEEEEEDDLSEASEAEOSEAEEVEESSEEDERENHNLIIVPESEPDSDSESEAEVEEGTPOSGALTEGDVPDPSPALDIETKOELPKYLPLALOGCSRVEPFOCLNIEGTGYGVYRKDKDTDEIVALTKRLMEKEKEGPFTLSRLREINTLKROHPNTVAIRELVGSNMDKIITYIMNYVBHLKSMTMKOPLPLEGVETKLTIQLLRGVKHLHDWMLIHRLKTSNLILSHAGILVSPPPSPSQDGDPGPTHSRPYSVAG"	

BASE COUNT	634	a	605	c	738	g	352	t
ORIGIN								

Query Match	100.0%	Score 222:	DB 9:	Length 2329:
Best Local Similarity	100.0%;	Pred. No. 1..1e-35;	Mismatches 0;	Indels 0; Gaps 0.
Matches 222;	Conservative	0;	Mismatches	0;


```

Qy      1  gacatcagcacagcgagagaagaaccacaactctgcgcgaagtctctgtcaagaagaatacagg 60
Db      689 GACATCAGCCACAGCGACAGAAGACCACTGCGCCGAGTCCTGTCACGACGAATCAGCG 748

Qy      61 tcaagttctcgaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 120
Db      749 TCAGGTTCTCAGGAAGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCACG 808

Qy      121 gaagaatcagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 180
Db      809 GAAGAATCAAGAGAGAGAAAGAGAGAGAGAGAGAGAGAGACC GGACACTCTGAAGAG 868

Qy      181 gcatcagcagcagtcctccgaagaagaagtaagtgaagaagaagt 222
Db      869 GCATCAGACGAGTCCTGCCGAAGAAGTAAGTCAGCAAGAAATG 910
  
```


RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE
HSU04818	2340 bp mRNA	Human protein kinase PITSLRE alpha 2-4 mRNA, complete cds.	U04818	004818.1	GI:507163	human.
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 2340) Xiang,J., Iahit,J.M., Grenet,J.A., Easton,J.B. and Kidd,V.J. Molecular cloning and expression of alternatively spliced PITSLRE protein kinase isoforms J. Biol. Chem. 269, 15786-15794 (1994)					
AUTHORS	JOURNAL MEDLINE 94253170 2 (bases 1 to 2340) Kidd,V.J. Direct Submission Submitted (03-JAN-1994) Vincent J. Kidd, St. Jude Children's					

Source

CORPORATE SOURCE: Y. Furuchi, AGENE Research Institute, 200 Kajiwara, Kamakura, Kanagawa 247, Japan. furuchi@gene.co.jp
 SOURCE: Molecular and Cellular Biology, (1998) 18/11 (6191-6200).
 Refs: 49
 ISSN: 0270-7306 CODEN: MCEBD4
 United States
 DOCUMENT TYPE: Journal; Article
 FILE SEGMENT: 029 Clinical Biochemistry
 LANGUAGE: English
 SUMMARY LANGUAGE: English

L6 ANSWER 4 OF 4 BIOSIS COPYRIGHT 2001 BIOSIS
 T1 **p53** Regulates the minimal promoter of the human topoisomerase II- α gene.
 AB DNA topoisomerase II- α is an essential enzyme for chromosome segregation during mitosis. Consistent with a cell division-specific role, the expression of the topoisomerase II- α gene is strongly influenced by the proliferation status of cells. The **p53** protein is one of the most important regulators of **cell cycle** progression in mammals, with an apparent dual role in the induction of **cell cycle** arrest following cytotoxic insults and in the regulation of the apoptotic cell death pathway. We have analysed whether **p53** plays a role in regulating expression of the human topoisomerase II- α gene. We show that wild-type, but not mutant, **p53** is able to decrease substantially the activity of the full length topoisomerase II- α gene promoter. Using a series of constructs comprising various deleted or mutated versions of the promoter lacking critical **cis-acting elements**, we show that this **p53**-specific regulation of the topoisomerase II- α promoter is independent of all characterized transcription factor binding sites and is directed at the minimal gene promoter. We conclude that expression of wild-type **p53** induces downregulation of the human topoisomerase II- α promoter by acting on the basal transcription machinery. These findings implicate topoisomerase II as one of the downstream targets for **p53**-dependent regulation of **cell cycle** progression in human cells.

ACCESSION NUMBER: 1997:19679 BIOSIS
 DOCUMENT NUMBER: PREV199799318882
 TITLE: **p53** Regulates the minimal promoter of the human topoisomerase II- α gene.
 AUTHOR(S): Sandri, M. Ines; Isaacs, Richard J.; Ongkeko, Weg M.; Harris, Adrian L.; Hickson, Ian D. (1); Brogini, Massimo; Vikhanskaya, Faina
 CORPORATE SOURCE: (1) Imperial Cancer Research Fund, Univ. Oxford, Inst. Molecular Med., John Radcliffe Hosp., Oxford OX3 9DU UK
 SOURCE: Nucleic Acids Research, (1996) Vol. 24, No. 22, pp. 4464-4470.
 ISSN: 0305-1048.
 DOCUMENT TYPE: Article
 LANGUAGE: English

=>

repression of the cdc2 promoter, while E1B-19K does not. Since the cdc2 promoter does not contain a TATA sequence, we performed deletion and point mutation analyses and identified the inverted CCAAT sequence located at -76 as a **cis-acting element** for the **p53**-mediated regulation. We found that a specific DNA-protein complex is formed at the CCAAT sequence and that this complex contains the NF-Y transcription factor. Consistently, a dominant negative mutant of the NF-YA subunit, NF-YA₂₉, decreases the cdc2 promoter, and **p53** does not further decrease the promoter activity in the presence of NF-YA₂₉. These results suggest that **p53** negatively regulates cdc2 transcription and that the NF-Y transcription factor is required for the **p53**-mediated regulation.

ACCESSION NUMBER: 2000:2169 BIOSIS
DOCUMENT NUMBER: PREV20000002169

TITLE: **p53** negatively regulates cdc2 transcription via

AUTHOR(S): Yun, Jeonho; Chae, Hee-Don; Choy, Hyon E.; Chung, Jongkyeong; Yoo, Hyang-Sook; Han, Moon-Hi; Shin, Deug Y.

CORPORATE SOURCE: (1) Bioscience Research Division, Korea Research Institute of Bioscience and Biotechnology, Yuseong, Taejeon, 305-600

SOURCE: Journal of Biological Chemistry, (Oct. 15, 1999) Vol. 274, No. 42, pp. 29677-29682.

ISSN: 0021-9258.

DOCUMENT TYPE: Article

LANGUAGE: English

SUMMARY LANGUAGE: English

L6 ANSWER 3 OF 4 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
TI **p53**-mediated transcription of the Werner Helicase gene is modulated by Rb

AB The regulation of Werner's syndrome gene (WRN) expression was studied by characterizing the cis-regulatory elements in the promoter region and the trans-activating factors that bind to them. First, we defined the transcription initiation sites and the sequence of the 5' upstream region (2.8 kb) of WRN that contains a number of cis-regulatory elements, including 7 Sp1, 9 retinoblastoma control element (RCE), and 14 AP2 motifs. A region consisting of nucleotides -67 to +160 was identified as the principal promoter of WRN by reporter gene assays in HeLa cells, using a series of WRN promoter-luciferase reporter (WRN-Luc) plasmids that contained the 5' - truncated or mutated WRN upstream regions. In particular, two Sp1 elements proximal to the transcription initiation site are indispensable for WRN promoter activity and bind specifically to Sp1 proteins. The RCE enhances WRN promoter activity. Coexpression of the WRN-Luc plasmids with various dosages of plasmids expressing Rb or **p53** in Saos2 cells lacking active Rb and **p53** proteins showed that the introduced Rb upregulates WRN promoter activity a maximum of 2.5-fold, while **p53** downregulates it a maximum of 7-fold, both dose dependently. Consistently, the overexpressed Rb and **p53** proteins also affected the endogenous WRN mRNA levels in Saos2 cells, resulting in an increase with Rb and a decrease with **p53**. These findings suggest that WRN expression, like that of other housekeeping genes, is directed mainly by the Sp1 transcriptional control system but is also further modulated by transcription factors, including Rb and **p53**, that are implicated in the **cell cycle**, cell senescence, and genomic instability.

ACCESSION NUMBER: 1998370638 EMBASE

TITLE: **p53**-mediated transcription of the Werner Helicase gene is

modulated by Rb and **p53**.

AUTHOR: Yamabe Y.; Shimamoto A.; Goto M.; Yokota J.; Sugawara M.; Furutachi Y.

L6 ANSWER 1 OF 4 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.DUPLICATE 1
 TI Repression of CDK1 and other genes with CDE and CHR promoter elements during DNA damage-induced G2/M arrest in human cells.
 AB Entry into mitosis is controlled by the cyclin-dependent kinase CDK1 and can be delayed in response to DNA damage. In some systems, such G2/M arrest has been shown to reflect the stabilization of inhibitory phosphorylation sites on CDK1. In human cells, full G2 arrest appears to involve additional mechanisms. We describe here the prolonged (>6 day) downregulation of CDK1 protein and mRNA levels following DNA damage in human cells. This silencing of gene expression is observed in primary human fibroblasts and in two cell lines with functional **p53** but not in HeLa cells, where **p53** is inactive. Silencing is accompanied by the accumulation of cells in G2, when CDK1 expression is normally maximal. The response is impaired by mutations in **cis-acting elements** (CDE and CHR) in the CDK1 promoter, indicating that silencing occurs at the transcriptional level. These elements have previously been implicated in the repression of transcription during G1 that is normally lifted as cells progress into S and G2. Interestingly, we find that other genes, including those for CDC25C, cyclin A2, cyclin B1, CENP-A, and topoisomerase II.alpha., that are normally expressed preferentially in G2 and whose promoter regions include putative CDE and CHR elements are also downregulated in response to DNA damage. These data, together with those of other groups, support the existence of a **p53**-dependent, DNA damage-activated pathway leading to CHR- and CDE-mediated transcriptional repression of various G2-specific genes. This pathway may be required for sustained periods of G2 arrest following DNA damage.
 ACCESSION NUMBER: 2000101579 EMBASE
 TITLE: Repression of CDK1 and other genes with CDE and CHR promoter elements during DNA damage-induced G2/M arrest in human cells.
 AUTHOR: Badie C.; Itzhaki J.E.; Sullivan M.J.; Carpenter A.J.; Porter A.C.G.; Porter A.C.G. Porter, MRC Clinical Sciences Centre, Imperial College School of Medicine, Hammersmith Hospital, London W12 0NN, United Kingdom. andy.porter@csc.mrc.ac.uk
 SOURCE: Molecular and Cellular Biology, (2000) 20/7 (2358-2366). Refs: 65
 ISSN: 0270-7306 CODEN: MCEBD4
 COUNTRY: United States
 DOCUMENT TYPE: Journal; Article
 FILE SEGMENT: 029
 LANGUAGE: English
 SUMMARY LANGUAGE: English
 L6 ANSWER 2 OF 4 BIOSIS COPYRIGHT 2001 BIOSIS
 TI **p53** negatively regulates cdc2 transcription via the CCAAT-binding NF-Y transcription factor.
 AB The **p53** tumor suppressor protein regulates the transcription of regulatory genes involved in **cell cycle** arrest and apoptosis. We have reported previously that inducible expression of the **p53** gene leads to the **cell cycle** arrest both at G1 and G2/M in association with induction of p21 and reduction of mitotic cyclins (cyclin A and B) and cdc2 mRNA. In this study, we investigated the mechanism by which **p53** regulates transcription of the cdc2 gene. Transient transfection analysis showed that wild type **p53** represses whereas various dominant negative mutants of **p53** increase cdc2 transcription. The cdc2 promoter activity is not repressed in cells transfected with a transactivation mutant, p5322/23. An adenovirus oncoprotein, E1B-55K inhibits the **p53**-mediated

(FILE 'HOME' ENTERED AT 09:05:12 ON 05 DEC 2001)

FILE 'BIOSIS', EMBASE, MEDLINE' ENTERED AT 09:05:39 ON 05 DEC 2001
L1 143956 S CELL CYCLE
L2 71863 S P53 OR CMYC
L3 11807 S L1 AND L2
L4 0 S L3 AND (INTERNAL RIBOSOME ENTRY SEQUENCE? OR IRES)
L5 9 S L3 AND (CIS-ACTING ELEMENT?)
L6 4 DUF REM L5 (5 DUPLICATES REMOVED)

block of induced erythroid differentiation. In this paper we report the effects of aromatic poly amides (TAPP-H, TAPP-B and DAPP-H) and their bromo-derivatives (TAPP-Br, TAPP-Br and DAPP-Br) on cell growth and erythroid differentiation of murine Friend erythroleukemic cells. These compounds are strong inhibitors of serine **proteases**. The results obtained give evidence (a) for inhibition of cell proliferation by all the aromatic poly-amides studied; (b) for stronger antiproliferative activity of the Br-derivatives; (c) for a lack of inhibition of globin mRNA accumulation and hemoglobin synthesis in uninduced cells as well as in Friend cells induced to erythroid differentiation by hexamethylenesuccinimide. These results suggest that this class of antiproliferative compounds exhibits differential effects on **cell-cycle** and differentiation specific genes. Thus, these aromatic poly-amides could be used in experimental therapy without interfering with differentiative processes also when combined with differentiating agents.

ACCESSION NUMBER: 89203563 EMBASE

DOCUMENT NUMBER: 1989203563

TITLE:

Friend erythroleukemic cells treated with aromatic polyamides exhibiting antiprotease activity: Inhibition of cell growth is not associated with block of induced erythroid differentiation.

AUTHOR:

Nastruzzi C.; Ferioto G.; Barbieri R.; Gambati R. Dipartimento di Scienze Farmaceutiche, Università di

SOURCE:

Cancer Journal, (1989) 2/10 (335-340).
ISSN: 0765-7846 CODEN: CANJFI

COUNTRY:

France

DOCUMENT TYPE:

Journal

FILE SEGMENT:

016

Cancer

025

Hematology

029

Clinical Biochemistry

037

Drug Literature Index

English

French; English

LANGUAGE:

SUMMARY LANGUAGE:

L15 ANSWER 19 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
TI Friend erythroleukemic cells treated with aromatic polyamides exhibiting
antiprotease activity: Inhibition of cell growth is not associated with

LANGUAGE: English
SUMMARY LANGUAGE: English; French
FILE SEGMENT: 016
DOCUMENT TYPE: 005
COUNTRY: France
SOURCE: Cancer Journal, (1993) 6/6 (322-329).
ISSN: 0765-7846 CODEN: CANJFI
CORPORATE SOURCE: Department of Pathology, Université Laval, L'Hotel-Dieu, 11
AUTHOR: Tetu B.; Brissson J.; Roberge N.
TITLE: Prognostic significance of tumor DNA content and S-phase
fraction by flow cytometry in node-positive breast
carcinomas.

AB Purpose - This study was aimed at evaluating the prognostic value of tumor
cytometry in node-positive breast carcinomas.
TI Prognostic significance of tumor DNA content and S-phase fraction by flow
L15 ANSWER 18 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
LANGUAGE: English
SUMMARY LANGUAGE: English
FILE SEGMENT: 025
DOCUMENT TYPE: 022
COUNTRY: France
SOURCE: Cancer Journal, (1995) 8/4 (195-200).
ISSN: 0765-7846 CODEN: CANJFI
CORPORATE SOURCE: Oncology Research Program, Toronto General Hospital, 67
AUTHOR: Guin B.; Padua R.A.
TITLE: Progression of leukemia?
AB Purpose - This study was aimed at evaluating the prognostic value of tumor
cytometry in node-positive breast carcinomas.
TI Prognostic significance of tumor DNA content and S-phase fraction by flow
L15 ANSWER 17 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
LANGUAGE: English
SUMMARY LANGUAGE: English
FILE SEGMENT: 016
DOCUMENT TYPE: 005
COUNTRY: France
SOURCE: Cancer Journal, (1993) 6/6 (322-329).
ISSN: 0765-7846 CODEN: CANJFI
CORPORATE SOURCE: Department of Pathology, Université Laval, L'Hotel-Dieu, 11
AUTHOR: Tetu B.; Brissson J.; Roberge N.
TITLE: Prognostic significance of tumor DNA content and S-phase
fraction by flow cytometry in node-positive breast
carcinomas.

AB Purpose - This study was aimed at evaluating the prognostic value of tumor
cytometry in node-positive breast carcinomas.
TI Prognostic significance of tumor DNA content and S-phase fraction by flow
L15 ANSWER 16 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
LANGUAGE: English
SUMMARY LANGUAGE: English
FILE SEGMENT: 016
DOCUMENT TYPE: 005
COUNTRY: France
SOURCE: Cancer Journal, (1993) 6/6 (322-329).
ISSN: 0765-7846 CODEN: CANJFI
CORPORATE SOURCE: Department of Pathology, Université Laval, L'Hotel-Dieu, 11
AUTHOR: Tetu B.; Brissson J.; Roberge N.
TITLE: Prognostic significance of tumor DNA content and S-phase
fraction by flow cytometry in node-positive breast
carcinomas.

AMERICAN, (1997 Jan-Feb) 3 (1) 13-20.

Journal code: CR8; 9513568. ISSN: 1081-4442.

United States

Journal; Article; (JOURNAL ARTICLE)

English

Priority Journals

199703

Entered STN: 19970407

Last Updated on STN: 19980206

Entered Medline: 19970327

AB

To summarize the majority of p53 mutations in leukemias are point

leukemias?

TI Do p53 mutations have a role in the initiation and progression of

leukemias? AB To summarize the majority of p53 mutations in leukemias are point mutations mainly in conserved regions IV and V, although CML patients also exhibit a notable portion of p53 gene rearrangements as seen more frequently in solid tumours. A significant proportion of the point mutations of the p53 gene occur at CpG dinucleotides, sites of DNA methylation which are believed to be preferentially mutated due to methylation-induced deamination of 5'-methylcytosine, perhaps further explaining the localization of p53 mutations to more specific sites within the functional domains of the gene in many cancers, including leukaemia. Arginine also appears to be a preferential site of mutations (at codon 175, 213, 248, 273 and 282). Arginine is coded for by CGN and hence a site for methylation, explaining part of its vulnerability to mutations. Disregarding these CpG preferential sites of mutation, the remaining mutations appear to be random in occurrence with a slight preference to G or C basepair mutations in MDS and lymphoma patients which may reflect the etiology of these diseases. DNA lesions occurring in p53 in hematological malignancies and solid tumours appear to follow a similar pattern of localization to the central DNA binding region of the p53 gene. Overall, the frequency of p53 mutations appears to increase with disease progression in all neoplasms. A higher rate of p53 mutations are found in AML patients than in the preceding MDS and in CML p53 mutations are more common in the blast crisis phase than in the preceding chronic phase or intermediate accelerated phase. This has also been found to be the case in solid tumours with an increasing frequency of p53 mutations with progression and/or metastatic propensity. The rate of p53 mutations has also been found to be increased in cell lines compared to their parent tumours, perhaps indicative of the growth advantage provided by such mutations and of the increased genomic instability imposed on tumours by p53 mutations. The p53 **cell cycle** checkpoint control for genetic integrity when mutated appears to allow molecular evolution with increased rapidly and cell division at a increased rate in the absence of the p53 G1/S stop signal. Mutations of the p53 gene are also increased in therapy-related leukemias compared to de-novo leukemias reiterating the DNA-damaging effects of many of the cancer treatment regimens. The frequency of p53 mutations in de-novo leukemias are low compared to most other malignancies. The exception to this appears to be CML, where mutations are reported to occur in as many as 30% of patients at blast crisis, some of the more severe forms of ALL and established leukemic cell lines. It is possible that another mechanism of p53 dysfunction, such as elevated MDM-2 expression may be the cause of the aberration of normal p53 function in leukemias. Another mechanism of cell deregulation unrelated to the p53 pathway may otherwise be involved in the development of hematological neoplasias. It is surprising that one gene has had such a central role in so many critical processes and hence in the development of such a diverse range of cancers.

ACCESSION NUMBER: 95261853 EMBASE

DOCUMENT NUMBER: 1995261853

TITLE: Do p53 mutations have a role in the initiation and

115 ANSWER 15 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
 TI Tumor-related prognostic factors for breast cancer.
 AB Interest in prognostic factors for breast cancer has been stimulated by the success of systemic adjuvant therapy for early-stage operable disease. Patients destined for recurrence can be selected for systemic adjuvant therapy, while patients not likely to recur can be spared the morbidity of unnecessary treatment. The number of tumor-related features available for prognosis has grown impressively in recent years. The purpose of this article is to review tumor-related biologic factors and relate them to prognosis and treatment objectives.
 ACCESSION NUMBER: 97034636 EMBASE
 DOCUMENT NUMBER: 1997034636
 TITLE: Tumor-related prognostic factors for breast cancer.
 AUTHOR: Donegan W.L.
 CORPORATE SOURCE: Dr. W.L. Donegan, Department of Surgery, Sinai Samaritan Medical Center, Milwaukee, WI, United States
 SOURCE: *Ca-A Cancer Journal for Clinicians*, (1997) 47/1 (28-51).
 Refs: 174
 ISSN: 0007-9235 CODEN: CAMCAM
 COUNTRY: United States
 DOCUMENT TYPE: Journal; General Review
 FILE SEGMENT: 005 General Pathology and Pathological Anatomy
 010 Obstetrics and Gynecology
 016 Cancer
 017 Public Health, Social Medicine and Epidemiology
 LANGUAGE: English
 SUMMARY LANGUAGE: English
 115 ANSWER 16 OF 19 MEDLINE
 TI Inhibition of Raf-1 **protein** kinase by antisense phosphorothioate oligodeoxynucleotide is associated with sensitization of human laryngeal squamous carcinoma cells to gamma radiation.
 AB BACKGROUND: Previous studies suggest a central role for Raf-1 **protein** kinase in mitogenic and radiation damage responsive signaling pathways. PURPOSE: Here we report that the expression and enzymatic activity of Raf-1 **protein** are inhibited in SQ-20B cells exposed to rat antisense oligodeoxynucleotide (As-ODNs) directed against the translation initiation site of human c-raf-1 cDNA. In contrast, treatment of SQ-20B cells with an equimolar concentration of rat sense oligodeoxy-ribonucleotide (S-ODNs) had no effect on the expression and activity of Raf-1. RESULTS: We have observed radiosensitization of rat As-ODNs-treated SQ-20B cells. The dose modifying factor of As-ODNs treatment was approximately 1.4. CONCLUSIONS: These studies demonstrate that rat As-ODNs is a DNA sequence-specific radiosensitizer that may have potential for use in the radiation therapy of cancers.
 ACCESSION NUMBER: 97205637 MEDLINE
 DOCUMENT NUMBER: 97205637 Pubmed ID: 9072302
 TITLE: Inhibition of Raf-1 **protein** kinase by antisense phosphorothioate oligodeoxynucleotide is associated with sensitization of human laryngeal squamous carcinoma cells to gamma radiation.
 AUTHOR: Soldatenkov V A; Dritschilo A; Wang F H; Olah Z; Anderson W B; Kasid U
 CORPORATE SOURCE: Department of Radiation Medicine, Lombardi Cancer Center, Georgetown University Medical Center, Washington, D.C.
 CONTRACT NUMBER: CA46641 (NCI)
 CA58984 (NCI)
 CA65012 (NCI)
 +
 SOURCE: CANCER JOURNAL FROM SCIENTIFIC

TI Altered regulation of specific **cell-cycle** genes as a marker for transformed and fetal cells.

ACCESSION NUMBER: 97278885 EMBASE
DOCUMENT NUMBER: 1997278885

TITLE: Altered regulation of specific **cell-cycle** genes as a marker for transformed and fetal cells.

AUTHOR: Hengstschlager M.; Soucek T.; Fusch O.; Braun K.; Bernaschek G.; Hengstschlager-Ottner E. M. Hengstschlager, Obstetrics and Gynecology, University of Vienna, Dept. of Prenatal Diagnosis/Ther., Wahringgasse 18-20, A-1090 Vienna, Austria

SOURCE: **Cancer Journal**, (1997) 10/4 (207-210). Refs: 20
ISSN: 0765-7846 CODEN: CANJFI
COUNTRY: France
DOCUMENT TYPE: Journal; General Review
FILE SEGMENT: 016 Cancer
LANGUAGE: English

LI5 ANSWER 14 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
TI Cell division cycle alterations in human malignancies.
AB This paper briefly reviews studies on the genes that control the cell division process in human carcinogenesis. It describes the key biochemical events of the division cycle, their regulation, and genetic alterations that occur in human malignancies. The process of cell division is the result of a cycling and time-regulated activation of specific enzymes that phosphorylate (and thus control) **proteins** required for progression toward cell mitosis. These enzymes, named cyclin-dependent kinases (CDKs), are activated by the binding of **protein** co-factors (i.e. cyclins) and are inhibited by specific **proteins** (CDK-inhibitors or CDIs). A large body of evidence demonstrates that several genes involved in **cell cycle** control are altered in human tumors. These include, in addition to the two archetypal tumor suppressor genes RB and p53, a CDK gene (CDK4), a cyclin gene (CCDN1) and a CDI gene (p16INK4A). In particular, p16(INK4A) gene is inactivated in many different malignancies with sometimes a greater than 80% incidence. These observations, along with studies carried out in animal models, suggest that almost all human cancers show alterations of the molecular mechanisms that control the division cycle. Such alterations might allow somatic cells to escape physiological cellular differentiation and senescence, thus favouring the development of further genetic aberrations and the progression towards an invasive malignant phenotype. In conclusion, we shall envisage possible scenarios for future clinical research (including potential diagnostic developments and new therapeutic strategies) that are related to the above biochemical investigations.

ACCESSION NUMBER: 97243439 EMBASE
DOCUMENT NUMBER: 1997243439

TITLE: Cell division cycle alterations in human malignancies.
AUTHOR: Della Ragione F.; Borriello A.; Giordani L.; Tolascio A. F. Della Ragione, Inst. of Biochem. of Macromolecules, Medical School, Second University of Naples, Via Costantinopoli 16, 80138 Napoli, Italy

SOURCE: **Cancer Journal**, (1997) 10/3 (151-156). Refs: 20
ISSN: 0765-7846 CODEN: CANJFI
COUNTRY: France
DOCUMENT TYPE: Journal; General Review
FILE SEGMENT: 016 Cancer
LANGUAGE: English

DOCUMENT TYPE: Journal; General Review
 FILE SEGMENT: 005 General Pathology and Pathological Anatomy
 LANGUAGE: English
 SUMMARY LANGUAGE: English
 L15 ANSWER 10 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
 TI p53 Expression, proliferative activity and prognosis in cancer.
 ACCESSION NUMBER: 1998376320 EMBASE
 TITLE: p53 Expression, proliferative activity and prognosis in cancer.
 AUTHOR: Pich A.
 AUTHOR: A. Pich, Dept. of Biomed. Sci./Human Oncology, University
 CORPORATE SOURCE: of Turin, Via Santena 7, 10126 Torino, Italy.
 SOURCE: Cancer Journal, (1998) 11/5 (223-228).
 Refs: 28
 ISSN: 0765-7846 CODEN: CANJFI
 FRANCE
 DOCUMENT TYPE: Journal; General Review
 FILE SEGMENT: 005 General Pathology and Pathological Anatomy
 LANGUAGE: English
 L15 ANSWER 11 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
 TI Dietary restriction, amino acid availability and cancer.
 ACCESSION NUMBER: 1998299937 EMBASE
 TITLE: Dietary restriction, amino acid availability and cancer.
 AUTHOR: Wheatley D.N.
 AUTHOR: D.N. Wheatley, Cell Pathology Unit, University Medical
 CORPORATE SOURCE: School, Foresterhill, Aberdeen AB25 2ZD, United Kingdom.
 SOURCE: Cancer Journal, (1998) 11/4 (183-189).
 Refs: 80
 ISSN: 0765-7846 CODEN: CANJFI
 FRANCE
 DOCUMENT TYPE: Journal; General Review
 FILE SEGMENT: 016 Cancer
 LANGUAGE: English
 L15 ANSWER 12 OF 19 BIOSIS COPYRIGHT 2001 BIOSIS
 TI Control of eukaryotic cell cycle progression by
 ACCESSION NUMBER: 1998:299160 BIOSIS
 DOCUMENT NUMBER: PREV199800299160
 TITLE: Control of eukaryotic cell cycle progression by phosphorylation of cyclin-dependent kinases.
 AUTHOR(S): Morgan, David O. (1); Fisher, Robert P.; Espinoza, F.
 AUTHOR: Herman; Farrell, Allison; Nourse, Jamison; Chamberlin,
 CORPORATE SOURCE: Holly; Jin, Pei
 (1) Dep. Physiol, Univ. California, 513 Ramassus Ave., San
 Francisco, CA 94143 USA
 SOURCE: Cancer Journal from Scientific American, (May,
 1998) Vol. 4, No. SUPPL. 1, pp. S77-S83.
 ISSN: 1081-4442.
 DOCUMENT TYPE: General Review
 LANGUAGE: English
 L15 ANSWER 13 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.

cycle. In normal and cancerous cells, a balance between proapoptotic and antiapoptotic signals exists. **Protein kinase C** (PKC) is a cellular serine/threonine kinase with a central role in the mediation of mitogenic signals as well as the regulation of anti-apoptotic signals. Inhibition of PKC by a novel group of chemical agents (PKC inhibitors) can induce apoptosis in some malignant cell lines, act as differentiating agents, and enhance the effect of cytotoxic chemotherapy. Other kinase inhibitors are potent inhibitors of kinases involved in the control of **cell cycle** progression (cyclin-dependent kinases (cdks)). cdk inhibitors are able to induce **cell cycle** arrest in neoplastic cells and also act as enhancers of chemotherapy-induced apoptosis. The catalytic domain of different classes of kinases (PKC and cdk) share considerable homology. As a result, many kinase inhibitors that act by blocking the catalytic site are not highly specific and may act as inhibitors of both PKC and cdk. Preliminary studies point to potential applications for some of these PKC/cdk inhibitors, and current clinical trials are exploring the role these agents might have in cancer therapy. In this article, we discuss the rationale for the development of this novel class of agents and highlight those drugs, which have shown promise in clinical testing.

ACCESSION NUMBER: 2001:24657 BIOSIS
DOCUMENT NUMBER: PREV200100024657
TITLE: Cyclin-dependent kinase and **protein kinase C** inhibitors: A novel class of antineoplastic agents in clinical development.

AUTHOR(S): Kaubisch, Andreas; Schwartz, Gary K. (1)
CORPORATE SOURCE: (1) Division of Gastrointestinal Oncology, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY, 10021 USA
SOURCE: **Cancer Journal**, (July-August, 2000) Vol. 6, No. 4, pp. 192-212. print.
ISSN: 1528-9117.

DOCUMENT TYPE: Article
LANGUAGE: English
SUMMARY LANGUAGE: English

L15 ANSWER 6 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
TI Cancer Journal: Editorial.
ACCESSION NUMBER: 2000374876 EMBASE

TITLE: Cancer Journal: Editorial.
AUTHOR: Devita V.T. Jr.; Hellman S.; Rosenberg S.A.
SOURCE: **Cancer Journal**, (2000) 6/4 (191).
ISSN: 1528-9117 CODEN: CAJOCB

COUNTRY: United States
DOCUMENT TYPE: Journal; Editorial
FILE SEGMENT: 016 Cancer Pharmacology 030 Drug Literature Index 037

LANGUAGE: English

L15 ANSWER 7 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
TI Primary central nervous system tumors: Advances in knowledge and treatment.
AB The ability to diagnose, monitor, and treat CNS tumors has been improved by new imaging techniques such as positron emission tomography (PET) scanning and functional MR imaging, stereotactic surgery, delivery of radiotherapy with brachytherapy and radiosurgery, and novel methods for delivering chemotherapy. These innovations combined with the new information about tumor pathogenesis and behavior revealed by molecular research give hope that more specific treatments for malignant CNS tumors will be developed in the future.
ACCESSION NUMBER: 1998393029 EMBASE

reverse transcriptase polymerase chain reaction analysis revealed a markedly distinct molecular expression profile in the CL lines: overexpression of basic fibroblast growth factor, interleukin-6, interleukin-8, vascular endothelial growth factor, transforming growth factor-beta, epidermal growth factor receptor, caveolin, and bcl-2 messenger RNAs and marked down-regulation of E-cadherin, p-53, and pentaerythritol tetranitrate. CONCLUSIONS Early administration of hormonal therapy after failure of first line treatment is associated with a profound clonal selection of aggressive AI variants, such as CL-1 and CL-2 lines. These tumor lines, with their parental counterparts, can serve as valuable tools for studying the cellular and molecular mechanisms of CAP progression and metastasis under hormonal therapy. CL-1 and CL-2 offer a unique and reproducible model for the evaluation of drug sensitivity and for other therapeutic modalities for advanced prostate cancer.

ACCESSION NUMBER: 2001:26304 BIOSIS
DOCUMENT NUMBER: PREV200100026304
TITLE: Androgen deprivation induces selective outgrowth of aggressive hormone-refractory prostate cancer clones expressing distinct cellular and molecular properties not present in parental androgen-dependent cancer cells.

AUTHOR(S): Tso, Cho-Lea; McBride, William H.; Sun, Jirong; Patel, Belur; Tsui, Ke-Hung; Pak, Sun H.; Gittlitz, Barbara; Calliaw, Randy; van Ophoven, Arndt; Wu, Lily; deKernion, Jean; Belidegrun, Arie (1)
CORPORATE SOURCE: (1) Department of Urology, UCLA School of Medicine, CHS 66-118, Los Angeles, CA, 90095-1738 USA
SOURCE: *Cancer Journal*, (July August, 2000) Vol. 6, No. 4, pp. 220-233, print.
ISSN: 1528-9117.
DOCUMENT TYPE: Article
LANGUAGE: English
SUMMARY LANGUAGE: English

L15 ANSWER 4 OF 19 EMBASE COPYRIGHT 2001 ELSEVIER SCI. B.V.
TI Thymidylate synthase and the **cell cycle**: What should we believe?
ACCESSION NUMBER: 2000374879 EMBASE
TITLE: Thymidylate synthase and the **cell cycle**
AUTHOR: Dolnick B.J.
CORPORATE SOURCE: Dr. B.J. Dolnick, Dept. of Experimental Therapeutics, Roswell Park Cancer Institute, Elm and Carlton Streets, Buffalo, NY 14263-0001, United States
SOURCE: *Cancer Journal*, (2000) 6/4 (215-216).
Refs: 6
ISSN: 1528-9117 CODEN: CAJOCB
COUNTRY: United States
DOCUMENT TYPE: Journal; Note
FILE SEGMENT: 016 Cancer
030 Pharmacology
037 Drug Literature Index
LANGUAGE: English

L15 ANSWER 5 OF 19 BIOSIS COPYRIGHT 2001 BIOSIS
TI Cyclin-dependent kinase and **protein kinase C** inhibitors: A novel class of antineoplastic agents in clinical development.
AB Malignant cells have acquired adaptations, which give them a growth and survival advantage over normal cells. One effect of many of these adaptations is that many cancerous cells are less likely to undergo programmed cell death (apoptosis) and, moreover, are resistant to chemotherapy-induced apoptosis. Other features of neoplastic cells are the loss of regulated or orderly progression through the **cell**

arrest point for both cell lines was before the point at which Rb phosphorylation takes place, yet the confluent H630RI cells had threefold higher p21 than confluent H630WT cells. DISCUSSION These data suggest that the 5-FU-resistant H630RI cell lines arrest at a later point in G0/G1 and have a potentially greater capacity for proliferation.

ACCESSION NUMBER: 2001:26305 BIOSIS
DOCUMENT NUMBER: PREV200100026305
TITLE: cell cycle regulation of the G0/G1 transition in 5-fluorouracil-sensitive and -resistant human colon cancer cell lines.

AUTHOR(S): McGinn, Cornelius J.; Pestalozzi, Bernhard C.; Drake, James C.; Glennon, M. Clay; Kunugi, Keith; Otterson, G.; Allegra, Carmen J.; Johnston, Patrick G.; Kinsella, Timothy J. (1) (1) Department of Radiation Oncology, University Hospitals of Cleveland, 11100 Euclid Ave, LTR 6068, Cleveland, OH, 44106-6068 USA

CORPORATE SOURCE: Cancer Journal, (July August, 2000) Vol. 6, No. 4, pp. 234-242, print. ISSN: 1528-9117.

SOURCE: DOCUMENT TYPE: Article
LANGUAGE: English
SUMMARY LANGUAGE: English

L15 ANSWER 3 OF 19 BIOSIS COPYRIGHT 2001 BIOSIS
DUPLICATE 2

TI Androgen deprivation induces selective outgrowth of aggressive hormone-refractory prostate cancer clones expressing distinct cellular and molecular properties not present in parental androgen-dependent cancer cells.

AB PURPOSE The mechanism of progression of human prostate cancer (CaP) cells under androgen ablation therapy remains unclear. To study the alternative pathways of CaP cell growth under conditions of androgen deprivation, androgen-independent CaP variants were selected and expanded from an androgen-dependent CaP line via an in vitro androgen deprivation treatment. Cellular and molecular properties of these androgen-independent variants were characterized both in vitro and in vivo and compared with those of their parental androgen-dependent cells. METHODS Androgen deprivation treatment of an androgen-dependent CaP cell line, LNCaP, was carried out by replacing culture medium with RPMI 1640 medium plus 10% charcoal-stripped serum. Cells that survived through the androgen deprivation treatment were harvested and expanded in the androgen-deficient culture medium and were designated CL-1. The CL-1 cells were also recultured in androgen-containing medium and designated CL-2.

The growth (cell cycle analysis, 3H-thymidine incorporation assay, growth expansion, and colonization efficiency), expression of CaP-associated markers (semiquantitative reverse transcriptase polymerase chain reaction), interaction with endothelial and bone marrow stromal cells, sensitivity to anticancer agents and radiation (growth inhibition), and tumorigenicity of CL-1 and CL-2 cells were determined and compared with these characteristics in parental LNCaP cells. RESULTS CL-1 and CL-2 cells are fast-growing cells when compared with parental LNCaP cells. They were capable of potentiating the growth of endothelial and bone marrow stromal cells in co-culture experiments and acquired significant resistance to radiation and to anticancer cytotoxic agents (Taxol(R), paclitaxel, vinblastine, and etoposide). In contrast to the poorly tumorigenic parental LNCaP cells, CL-1 and CL-2 lines proved highly tumorigenic, exhibiting invasive and metastatic characteristics in intact and castrated mice or in female mice within a short period of 3 to 4 weeks. No growth supplements (e.g., Matrigel) were needed. When transfected with the green fluorescence protein (GFP) gene and transplanted orthotopically in the accessory sex gland, extensive metastatic disease from the primary CaP tumor could be identified in bone, lymph nodes, lung, liver, spleen, kidney, and brain. Semiquantitative

L15 ANSWER 1 OF 19 MEDLINE
 TI Signaling pathways in pancreatic cancer.
 AB Pancreatic cancer is, indisputably, one of the most malignant gastrointestinal tumors. Although the etiology of this disease is unknown, it is clearly linked to alterations in the biologic activities of various signaling molecules. Aberrant signaling activities of growth factors and their receptors, transcription factors, and proteins that control the cell cycle have been increasingly implicated in the pathogenesis and dissemination of pancreatic tumors. It is indeed possible that several of these molecules are, in fact, part of a signaling network that has gone awry. This review summarizes some recent advances in an attempt to generate a working model for future investigations.

ACCESSION NUMBER: 2001513706 IN-PROCESS
 DOCUMENT NUMBER: 21445507 PubMed ID: 11561604

TITLE: Signaling pathways in pancreatic cancer.
 AUTHOR: Reddy S A

CORPORATE SOURCE: Department of Gastrointestinal Medical Oncology, University of Texas M.D. Anderson Cancer Center, Houston 77030, USA.
 SOURCE: CANCER JOURNAL, (2001 Jul-Aug) 7 (4) 274-86.

PUB. COUNTRY: United States
 Journal code: DUN; 100931981. ISSN: 1528-9117.

LANGUAGE: English
 FILE SEGMENT: IN-PROCESS; NONINDEXED; Priority Journals
 ENTRY DATE: Entered STN: 20010920
 Last Updated on STN: 20010920

L15 ANSWER 2 OF 19 BIOSIS COPYRIGHT 2001 BIOSIS
 TI Cell cycle regulation of the G0/G1 transition in
 AB 5-fluorouracil-sensitive and -resistant human colon cancer cell lines. PURPOSE Resistance to 5-fluorouracil (5-FU) has been associated with thymidylate synthase (TS) gene amplification and increased TS protein levels. Increased TS protein expression has also been found to be a significant independent prognostic factor for disease-free survival and overall survival in patients treated with adjuvant 5-FU-based chemotherapy. In these studies and in our prior preclinical studies, TS has been considered a marker of proliferative capacity. The purpose of the current study was to further evaluate the association between TS levels and cell cycle kinetics in a 5-FU-resistant cell line with constitutive overexpression of TS. The influence of increased TS levels on cell cycle progression may provide insight into methods to overcome 5-FU resistance. MATERIALS 5-FU-sensitive NCI H630WT and 5-FU-resistant NCI H630R1 (with 15- to 20-fold higher TS protein levels) were utilized in this investigation to determine the influence of constitutive overexpression of TS on cell cycle kinetics. RESULTS There was no apparent influence of increased TS levels on cell cycle distribution during asynchronous growth, and both cell lines reach plateau growth phase in 120 hours, arresting in G0/G1, as determined by flow cytometry. In the H630WT cells, this G0/G1 arrest was associated with a 14- to 17-fold reduction in TS activity and protein levels (using the TS-106 monoclonal antibody), whereas in the H630R1 cells, only a two- to fivefold reduction was noted. Flow cytometry analysis utilizing Ki-67 indicated that there was no evidence of a G0 population in the confluent H630R1, whereas 26% +/- 7% of confluent H630WT cells were Ki-67 negative (G0) and the remainder had low Ki-67 signal intensity. Analysis of pRb phosphorylation and p16 and p21 expression suggested that the

FILE 'CANCERLIT' ENTERED AT 09:17:08 ON 05 DEC 2001	L7
2 S (CANCER JOURNAL)/CS	L8
0 S L7 AND 1998<=PY<=1998	L9
0 S CELL CYCLE DEPENDENT TRANSLATION	L10
0 S BIOSIS EMBASE MEDLINE CAPLUS CANCERLIT	
FILE 'BIOSIS, EMBASE, MEDLINE, CAPLUS, CANCERLIT' ENTERED AT 09:20:48 ON	
05 DEC 2001	L11
216006 S CELL CYCLE	L12
5674763 S PROTEIN?	L13
107329 S L11 AND L12	L14
31 S L13 AND CANCER JOURNAL/SO	L15
19 DUP REM L14 (12 DUPLICATES REMOVED)	